THE INTERACTION BETWEEN AUDITORY IMAGERY VIVIDNESS AND AUDITORY PERCEPTION, AND ITS APPLICATION TO HALLUCINATION PRONENESS

by

GEMMA RUTH GRAY

A thesis submitted to
The University of Birmingham
For the degree of
DOCTOR OF PHILOSOPHY

School of Psychology
College of Life Sciences
The University of Birmingham
September 2009

Acknowledgements

Firstly I'd like to thank Prof. Glyn Humphreys for his help throughout my research and for his patience and willingness to allow me room to work in my own way. I'd also like to express to Prof. Max Birchwood for his invaluable advice into the studies into hallucination proneness.

Many thanks to Ali Chowdhury for his help in programming the sequence for my fMRI study, and huge thanks to Pia Rotshtein for all her help with the design and analysis of this study. Thank you for your unwavering patience with me - I'd have been entirely lost without you!

I'd also like to thank Noreen O'Sullivan, Nele Demeyere and Sarah Houthuys for being wonderful office mates over the past four years and for all their support during the ups and downs of research – definitely couldn't have got through it without you guys! Thanks also to Christine Haecker, Giles Anderson, Hayley Wright and Lara Harris for all their kind words, sympathy and for being brilliant friends throughout the time I have known them. I'd also like to express my gratitude to my family for their love and belief in me over the years, and for helping me to accomplish everything that I've wanted to in life.

Finally, I'd like to thank my partner Tim for all his support, for keeping my feet firmly on the ground and for his insistence throughout the years that "it'll be fine".

Dedication

For my dear Nan and Granddad – yes, I'm still getting my sums right!

Abstract

Auditory imagery is commonly used in everyday life, yet the majority of imagery research has focused on the visual domain. This thesis determined some of the mediators of auditory imagery vividness and investigated how vividness affects the interaction between imagery and perception (Chapter 2). In addition, an fMRI study investigated the neural correlates of auditory imagery and perception (Chapter 3). The final empirical chapters assessed the interaction between auditory imagery vividness and hallucination proneness, and the influence of hallucination proneness on the interaction between imagery and perception (Chapters 4, 5 and 6). Imagery vividness differs according to sound category and familiarity and is affected by cues to imagine sounds. Imagery and perception can also interact to influence the detection of sounds in noise, and are processed by partially overlapping regions of the auditory cortex. Studies into hallucination proneness revealed little differences between high and low hallucination proneness participants when detecting sounds in noise. When stimuli had an emotional connotation (i.e. auditorily presented emotional words) high and low hallucination prone participants differed only in their memory recall rate, but not in their vividness ratings, or in their sound detection performance for such words.

Taken together, this thesis demonstrates that auditory imagery vividness is a robust measure that is affected by a range of cognitive factors. Vividness can influence detection of sounds in noise and has measurable affects on neural activation. These studies provide evidence for the theory that imagery and perception rely on overlapping areas of processing. The thesis also finds little association between hallucination proneness and auditory imagery vividness or sound detection performance. This suggests that factors other than auditory imagery are associated with proneness to hallucination-like experiences

Figures

Tables

Chapter One. Literature Review

In everyday life, auditory imagery frequently used, for instance when we try to recall a telephone number or when we replay a song in our heads. This domain of imagery has received relatively little research interest compared to other modalities however. The current thesis is an exploration of auditory imagery and perception, to investigate four main topics: (i) the factors that influence subjective experience of auditory imagery vividness, (ii) how auditory imagery and perception interact, (iii) the similarities in neural activation between imagery and perception, and (iv) the interaction between hallucination proneness and auditory imagery.

This chapter gives a background to previous imagery research. It reviews early research into mental imagery and studies examining the similarities in acoustic representations between imagery and perception for sounds, such as pitch, tempo and timbre. It also focuses on research into imagery vividness. This chapter presents evidence that imagery and perception for sounds can interact with each other to influence performance on sound detection tasks. This review also presents studies demonstrating a link between auditory imagery and experience of hallucination. Finally studies into the brain regions associated with auditory imagery and perception are presented. The neuroscientific studies reviewed demonstrate that imagery and perception activate similar regions of the brain and that there are dissociations in activation according to the sound category heard or imagined. In addition studies into auditory recognition disorders following brain injury give a further insight into the cortical organisation of the auditory system.

1. What is imagery?

Mental imagery involves the internal representation of perceptual processes in the absence of external stimulation e.g. “seeing with the mind′s eye” when using visual imagery (Kosslyn, Ganis, & Thompson, 2001). Such mental imagery has fascinated philosophers, psychologists and psychiatrists for centuries, as mental imagery is involved in many aspects of our everyday life - from remembering past events, to planning the future (Eardley & Pring, 2006). The term ‘mental imagery’ applies to all sensory modalities though the majority of studies have focused on the visual modality.

Due to its phenomenological nature imagery is difficult to assess directly, and, relative to studies of perceptual processes themselves, the study of imagery can be said to have been relatively neglected. One way to investigate imagery is to assess how performance on different tasks is affected by using imagery. For instance Paivio (1969) investigated the effects of imagery on memory performance, finding that verbal recall increased for words that were easier to create visual images for. This suggested a functional mnemonic role of imagery in memory processes. Considerable debate has ensued however, to determine the nature of imagery that may affect cognitive processing.

For instance Depictive Theory posits that information-processing in imagery relies on depictive images formed, like a screen that images are projected on, which can be analysed and assessed (Kosslyn, 1973). Kosslyn (1973) found that it took longer for participants to imagine moving between two features in imagined pictures (e.g. a map) depending on how far away they were in the actual picture. This suggested that participants were actually mentally scanning across the pictures to perform the task. Similarly Shepard and Metzler (1988) found that when participants mentally rotated imagined complex figures, it took longer for them to perform the rotation the further it was from the original orientation of the stimulus. These studies suggest that images reflect spatial relationships and orientations in a way similar to the processing of items that are actually perceived.

The opposing Propositional Theory posits that visual images are a consequence of the cognitive process, rather than the evidence of the process itself (i.e. "like the heat from a light bulb, when turned on to light the reading of a book", Kosslyn, 1994). Pylyshyn (1973) suggested that information-processing during imagery is similar to other kinds of cognition, and is based on a set of symbolic descriptions or propositions. Furthermore, Pylyshyn (1973) claimed that Kosslyn’s (1973) findings were likely to be due to experimenter or participant expectancies about the results or due to the participants’ “tacit knowledge” about the physical world, which is used when participants solve such imagery tasks. So for instance, because of the tacit knowledge that the time to scan between two points on an actual map varies according to the distance between them, participants may assume that it will take a similar amount of time when asked to perform the task using imagery. Pylyshyn (1973) found that imagery tasks where tacit knowledge cannot be used were not solved as well as those involving such knowledge, which was taken as evidence for Propositional Theory. On the other hand, further studies by Jolicoeur and Kosslyn (1985) determined that manipulation of either the experimenter’s or the participant’s bias did not affect the results of mental scanning experiments.

Recent advances in neuroimaging have marked a resurgence in interest in imagery. Such studies revealed areas in the brain responsive to visual perception, including the retinotopic cortex (Fox, Miezin, Allman, Vanessen, & Raichle, 1987; Engel, Glover, & Wandell, 1997) are activated during (visual) imagery tasks (DEsposito et al., 1997; O'Craven & Kanwisher, 2000; Wheeler, Petersen, & Buckner, 2000). These findings support the theory that visual images are depictive in nature.

2. Measurement of imagery

A key issue regarding imagery research is that imagery reflects phenomenological experience, and cannot be observed directly. Therefore accurate measurement of imagery abilities is difficult.

Imagery questionnaires are typically used to measure peoples’ self-reported ability to form mental images, and scores on these imagery scales are often related to some objective measure of memory. The Questionnaire upon Mental Imagery (QMI) is a commonly used measure of imagery vividness. The original questionnaire, designed by Betts in 1909, had over 100 items measuring different modalities of imagery (visual, auditory, motor, taste and olfaction). Participants rated how clearly they could imagine each item on the QMI using a 7-point scale. Sheehan (1967) created a shorter 35 item version, based on factor analysis of the original questionnaire, which is the most commonly used form to date. Use of this shorted form reveals high levels of inter-modality correlation, suggesting a general ‘vividness’ factor in imagery (Sheehan, 1967). Later studies have failed to find an association between the QMI and memory recall however, calling into question the validity of this questionnaire at measuring imagery processes and/or the role of imagery in memory (Richardson, 1979; Berger & Gaunitz, 1977).

Marks (1973) also developed a self-report imagery vividness questionnaire, called the Vividness of Visual Imagery Questionnaire (VVIQ). Like the QMI, the VVIQ required participants to rate the vividness of their imagery using a rating scale, for a range of items. Unlike the QMI however, the VVIQ focussed purely on visual imagery and contained several 'scenes' with each item increasing the complexity of the 'scene'. Marks (1973) then grouped participants as 'good visualisers' or 'bad visualisers' according to their average vividness ratings. These participants subsequently completed a memory task, in which they viewed 15 photographs, and after a delay were asked questions about those photographs. The study revealed that 'good visualisers' were significantly better at answering questions about the pictures. McKelvie & Demers (1979) later confirmed this finding in their study of short- and long-term memory recall. These studies therefore suggest that subjective vividness of visual imagery was a good predictor of memory abilities.

As mentioned above however other studies have questioned the validity of these measures, suggesting that vividness may be too idiosyncratic or a reflection of social desirable responding. Judging the vividness of one's own image inherently involves comparison with our perceived ideas of other peoples imagery and it is not clear how much this judgemental factor plays, compared to any perceptual factors. In addition there may be a bias towards giving social-desirable responses, which may contaminate ratings. For instance Allbut (2008) found an association between visual imagery vividness and measures of social-desirable responding.

More recently neuroimaging evidence has suggested an association between increased visual imagery vividness and stronger neural activation in the early visual cortex, implying that vividness of imagery does reflect increased strength of perceptual representations (Cui, Jeter, Yang, Montague, & Eagleman, 2007). In addition Belardinelli et al. (2009) found greater involvement of sensory-specific regions during imagery for items in different modalities, for high imager compared to low imagers.

2. Studies of auditory imagery

Though the visual modality has dominated imagery research in the past, interest is now increasing in auditory imagery. The methods used to assess auditory imagery are often similar to those used in visual imagery. Here the review will focus on three types of study into auditory imagery will be reviewed here, namely (i) individual differences in imagery abilities and (ii) behavioural similarities between auditory imagery and perception, (iii) neuroscientific and neuropsychological investigation into imagery and perception.

Vividness of Auditory Imagery

Studies of auditory imagery vividness are limited in number compared to studies in other modalities of imagery (Hubbard, 2010). Imagery questionnaire studies have revealed good correlation between imagery vividness in audition and imagery vividness in other modalities, however (Allbutt, Ling, Heffernan, & Shafiullah, 2008; Belardinelli et al., 2009; Sheehan, 1967).

Baddeley and Andrade (2000) went beyond the use of questionnaires by assessing how imagery vividness ratings varied for familiar and unfamiliar visual and auditory items, when participants either performed a spatial tapping secondary task or a counting secondary task. Auditory imagery vividness was lower following the counting task, and visual imagery vividness was lower following the spatial tapping task. This suggests phonological loop and visual-spatial sketchpad involvement in the respective imagery tasks. Further, vividness was higher for familiar items compared to unfamiliar items. This therefore suggests that interfering with specific components of working memory can differentially affect different modalities of imagery, and suggests an influence of long-term memory on vividness ratings. In addition memory for items used on the tasks showed the same pattern of results as that for imagery vividness, suggesting good association between imagery and memory.

As mentioned above, some studies question the validity of subjective vividness ratings. Allbutt et al. (2008) studied the association between auditory imagery vividness, visual imagery vividness and social-desirable responding. This study revealed an association between increased visual imagery vividness and increased "self-deceptive enhancement" (bias related to social desirability). Auditory imagery did not correlate with any measures of social-desirable responding, however. The authors suggested that having more vivid visual imagery may be perceived as more desirable than vivid auditory imagery, because of the perceived gains of having visual imagery, e.g. the ability to remember information more clearly and the importance of visualisation.

Similarities between imagery and perception

Research has focused on three main categories of sound: environmental sound, music and speech, and early studies focused on the acoustic similarities between imagery and perception for such sound categories.

Intons-Peterson et al. (1980; 1992) found that participants judged the loudness and pitch of imaged sounds similarly to when they actually heard the sounds. Similarly, music imagery and perception share temporal similarities. Halpern (1988) found that in both perception and imagery, participants took longer to say whether two song lyrics were the same pitch or not, dependent on how faraway the lyrics were from each other in the song. Crowder (1989) examined imagery for timbre (i.e. instrument voice). Participants were quicker at saying that two notes were the same, when the same instrument played both notes, rather than different instruments. In addition when participants imagined a particular instrument playing a note, they made "same" decisions quicker if the heard note and imagined note were played on the same instrument (i.e. if they imagine a trumpet playing a 'C' they were quicker to say 'same' if the heard note was a trumpet playing a 'C' rather than a flute playing a 'C'). Similarly, Zatorre et al. (1996) asked participants to compare pairs of instruments and say how similar they sound to each other, when imagined and when perceived. This revealed a similar matrix structure between imagery and perception, suggesting that participants can 'hear' the timbre similarities between sounds in imagery, similarly to perception. These studies have showed that auditory images have some affinity to auditory perceptions in terms of the physical characteristics of sounds.

Recent neuroimaging studies revealed that auditory imagery and perception rely on similar brain regions (see below for a review). Therefore performing one type of task may influence the other (Segal & Fusella, 1970).

Segal and Fusella (1970) examined the affect of visual and auditory imagery on visual or auditory target detection. In their first experiment, they asked participants to imagine familiar or unfamiliar visual or auditory item (e.g. seeing a volcano or hearing a typewriter) and then detect a faint visual or auditory target in noise (e.g. a light or a harmonica chord). Forming an image in the same modality to the target impaired detection sensitivity, compared to forming images in the other modality. This suggests that, though imagery may have a distracting affect on detection regardless of the modality, there is a modality-specific interference effect consistent with representational overlap between imagery and perception. Segal and Fusella (1970) interpreted this as imagery and perception competing for processing capacity, leading to detection impairments in perceptual tasks. These authors also found that imagining unfamiliar items hindered detection more than familiar items, which they attributed to unfamiliar images being more effortful to produce and therefore use more processing capacity.

However Farah and Smith (1983) reported opposite affects to that of Segal and Fusella (1970), finding a facilitatory effect of imagery on detection. They argued that imagery can aid detection if the image and the target match. Farah and Smith asked participants to image one of two frequencies of tone and then either detect the same frequency, or a different frequency tone in white noise. Detection of the tone improved if it matched the imagined tone. Farah and Smith (1983) suggested that the image of the target allowed allocation of attentional resources to the target item, making it easier to detect the target. The exact nature of the representational overlap may be critical to patterns of facilitation or interference.

Studies have also found that listening to a sound influences imagery tasks. Tinti, Cornoldi, and Marschark (1997) examined the affect of visual and auditory detection tasks on visual and auditory imagery tasks. Participants formed interactive auditory or visual images to pairs of words (e.g. if given ‘drill’ and ‘helicopter’ they could imagine the two sounds occurring together, or a visual image of a drill piercing the helicopter). Following each pair, participants detected either a visual or an auditory target in noise. At the end of the task, participants were given the first word of each pair, and tried to recall the second word. Concurrent detection of auditory compared to visual targets selectively impaired auditory image recall, and vice versa for visual images – mirroring the results from Segal and Fusella (1970). The data again show that perception and imagery interact, perhaps due to sharing of their neural representations.

Neuroscientific and neuropsychological studies of imagery and perception

3. Neuroanatomy of auditory system

Early studies of auditory imagery for different sound types suggest that imagery shares many characteristics with perception of those sounds. In the past, difficulties in imaging the regions involved in perception and imagery of sounds and the confounding effect of scanner noise has complicated investigation of the neural correlates of such processes (e.g. PET and fMRI, Boatman, 2006). Recent advances with these techniques has led to an increase in auditory investigations, however (Boatman, 2006).

A brief description of the peripheral and central auditory areas is necessary, before review of neuroimaging studies however. The peripheral auditory area includes the outer, middle and inner ears. On presentation of a sound, sound waves are channelled through the outer ear, down the ear cannel to the tympanic membrane. These sound waves cause the membrane to vibrate and move three middle ear bones, which in turn move the basilar membrane. The basilar membrane contains tonotopically (pitch) organised hair cells, which selectively move in response to specific sound frequencies. Action potentials (generated by movement of these cells) are transmitted down auditory nerve fibres in to the central auditory areas. These nerve fibres first transmit to the cochlear nuclei in each hemisphere, and each cochlear nucleus receive ipsilateral projections (i.e. input from the left auditory nerve fibres only goes to the left cochlear nucleus and the same for the right auditory nerve fibres). From here, connections lead to both the contralateral and ipsilateral superior olives and then to the inferior colliculus. Different auditory functions involve this region, such as frequency analysis, sound localisation, filtering and integration (Boatman, 2006). From this region, projections are sent from contralateral and ipsilateral sides to the medial geniculate body (MGB) in the thalamus, and then to the Primary Auditory Cortex.

The key structures in this area are the Heschl's gyrus and the planum temporale. The Heschl's gyrus is on the upper side of the superior temporal gyrus (STG) and in the Sylvian fissure comprises about two thirds of the auditory cortex. The planum temporale is located anteriorly and posteriorly to Heschl's gyrus. Surrounding this area is the Association Auditory Cortex, otherwise known as the auditory belt area (Kaas & Hackett, 2008). These regions are partially tonotopic and involved in the identification, localisation and discrimination of various types of sound. In addition, hemispheric lateralisation differs according to the task demands, the acoustic properties and the types of stimulus (Jensen, 2005), some of which are discussed below.

4. Neuroscientific studies: perception and imagery

Recent studies examine whether imagery for sound, without sound stimulation, activates auditory areas. Using fMRI, Yoo et al. (2001) found primary and secondary auditory cortex activation during auditory imagery. These regions are also active during auditory perception (Specht & Reul, 2003). Most studies exploring the similarities between imagery and perception focus on more complex stimuli, such as music and speech however.

4.1 Music

Halpern, Zatorre, Bouffard and Johnson (2004) investigated the association between behavioural and neural responses during an imagined and perceived musical timbre task. Behavioural investigation showed good association between imagery and perception (see also Zatorre et al., 1996, above). Auditory perception activated the right primary and secondary auditory areas, while the imagery task activated secondary auditory areas and the supplementary motor area, suggesting the use of subvocalisation.

The data from Halpern et al. (2004) contrast with those of Kraemer, Macrae, Green, and Kelley (2005) who found activation in the left auditory association cortex during silent gaps in familiar songs compared to unfamiliar songs. Further this effect held across songs with and without lyrics, suggesting the verbal content of the songs did not explain this asymmetry. These differences may reflect the difference in task demands: Halpern et al.'s task involved active timbre comparison, while Kraemer et al.'s study involved passively listening to melodies.

4.2 Speech

Speech perception is classically associated with activation in the left auditory areas, so that (for example) patients with lesions to the left temporal lobe have disorders in perceiving and understanding speech (Boatman, 2006). Neuroimaging studies corroborate this hemispheric lateralisation. Using fMRI, Binder et al. (1996) found considerable left lateralised activation (particularly in the planum temporale and superior temporal sulcus (STS)) when participants passively listened to monosyllabic nouns. Later studies support the finding of the left hemisphere activation in response to speech perception (Scott, Blank, Rosen, & Wise, 2000; Dick et al., 2007). McGuire et al.'s (1996a) PET studies revealed an association between inner speech and left inferior frontal gyrus activation. This study also revealed an association between auditory verbal imagery and extensive left hemisphere activation, including the premotor cortex, SMA, and temporal cortex. Similarly Shergill et al.'s (2001) fMRI study revealed an activation of the left inferior frontal and temporo-parietal cortices when participants imagined themselves speaking (inner speech).

Recent studies have criticised the theory of a left lateralised network however, with the argument suggesting more bilateral or task-related activation during speech processing (Binder et al., 2000; Price, Thierry, & Griffiths, 2005; Specht et al., 2003).

4.3 Environmental Sounds

Hoshiyama, Gunji and Kakigi's (2001) MEG study investigated activation during silent and sound-appropriate videos of a hammer hitting an anvil. During the silent videos, participants imagined the sound associated with the scene, revealing activation in the inferior sulcus and insula of the right hemisphere. This is consistent with right hemisphere specialisation for non-verbal sounds (Boatman, 2006), but the study only examined activation to a hammer sound and is therefore unlikely to be representative of all environmental sound processing.

Bunzeck, Wuestenberg, Lutz, Heinze, and Jancke (2005) study involved a larger range of items. During fMRI scanning participants saw videos of common everyday objects, with sound (perception condition) or without sound (imagery condition). The control condition involved presentation of scrambled videos. This revealed bilateral primary and secondary auditory cortex activation during sound perception, and bilateral secondary (but not primary) auditory cortex activation during imagery. Therefore the theory of right hemisphere lateralisation for non-verbal sounds may not be as clear-cut as previously suggested.

5. Dissociations between sound categories

Brain imaging studies suggest that imagery and perception within a sound category result in overlap in neural activation. Many other studies explored dissociations between different sound categories, for instance between speech and non-speech sounds (i.e. music, animal sound, environmental sounds) as these vary greatly in their acoustic and semantic characteristics (Binder et al., 2000; Boatman, 2006; Cummings et al., 2006; Dick et al., 2007; Specht et al., 2003). Specht and Raul (2003) explored activation for tones, non-verbal sounds (music and animal sounds) and speech using fMRI, revealing activation in bilateral primary and secondary auditory areas. The strength of this activation varied according to sound category however. Pure tones resulted in weak activation in the bilateral superior temporal sulci, whilst music and animal sounds resulted in stronger bilateral activation (mainly on the right-hand side) in the superior temporal sulci. Speech resulted in strongest activation, mainly in the left superior temporal sulci and the left inferior frontal gyrus. In this experiment the 'non-verbal sounds' included music and animal sounds, despite these two sound categories having different acoustic characteristics. Different results may have been found had the two sources of sound been split into two categories (Kraut et al., 2006).

The category of 'non-speech' sounds is in itself extremely broad, encompassing sounds of vehicles, tools, animals and music. In addition the environmental sound category incorporates both animal and non-living sounds. Animal sounds tend to have more harmonic content than other environmental sounds, however (Lewis, Brefczynski, Phinney, Janik, & Deyoe, 2005), and often resemble human vocalisations in their spectral content (Boatman, 2006). Consequently, recent studies investigated category-specific differences within the category of environmental sounds to explore whether dissociations exist.

In Lewis, Brefczynski, Phinney, Janik and Deyoe (2005) fMRI study, participants listened to various animal and tool sounds. Animal sounds resulted in bilateral medial superior temporal gyrus activity (mSTG), and widespread left hemisphere activation occurred during tool sound perception, including the motor cortex and a mirror neuron network. Lewis et al. (2005) suggested that these latter activations were due to the sounds corresponding to hand-manipulated tools, which result in greater multimodal response compared to other sounds. Altmann et al.(2007) also found bilateral STG activation in response to animal sounds, and Doehrmann et al. (2008) found selective adaptation of the left Heschl's gyrus animal vocalisations, and the left posterior middle temporal gyrus (pMTG) to tools sounds. These studies support the theory that animal and tool sounds form distinct semantic categories within the brain.

So far no studies have compared imagery for animal sounds compared to non-living environmental sounds, although Kraut et al. (2006) explored memory for animal and tool sounds. In this fMRI study, participants listened to animal and environmental sounds and judged whether the sounds belonged to real stimuli. Activation increased in the right STG, the left MTG, and left frontal lobe for animal sounds compared to environmental sounds. Kraut et al. (2006) suggested that such dissociations may suggest the existence of category-specific processing regions in the auditory semantic system that group sounds made by living things, similar to that found in the lexical-semantic and visual system. The sensory-functional hypothesis is often used to describe category-specific differences (Warrington and McCarthy, 1983). This posits that animals can be distinguished from each other by their sensory characteristics, whereas objects are usually distinguished from each other by their functional characteristics (see Humphreys and Forde, 2001). This theory can account for the double dissociations in category specific deficits for living and non-living things that patients with semantic deficits have. Interestingly, recent studies found similar category specific deficits in patients with relatively specific auditory agnosia (Kolinsky et al., 2002), suggesting similar organisation for the semantic systems of auditory sounds and visual objects.

Wu, Mai, Chan, Zheng, and Luo (2006) investigated imagination for animal sounds, in isolation of other non-living environmental sounds, using ERP. Participants viewed pictures of animals and either just looked at the pictures, or looked at them while imagining the sound the animal makes. This revealed significant differences in the P2 component, which the Wu et al. (2006) attribute to attention allocation, and a further difference between the two conditions emerging between 350 – 600ms post stimulus, attributed to imagery use. Unfortunately Wu et al. (2006) did not assess the similarities in the pattern of responses for imagery and perception.

This collection of studies highlights the difference that exists in activation, associated with perception of different categories of sounds. There have been few attempts, to factor-out the acoustic and semantic differences that may contribute to dissociations in activation however.

6. Patient studies of imagery and perception

Patient studies provide further evidence for the dissociation in cortical structures processing different sounds. Agnosia is an impairment in object recognition from the senses, without significant impairment to the sensory organs. Visual agnosia cases are perhaps the most common, including reports of patients with selective impairments for different types of objects, such as living things, non-living things, tools, musical instruments etc. (Warrington & Shallice, 1984). In addition, Lissauer (1988) defined two main subtypes of agnosia, based on whether the impairment affects perceptive information about objects (apperceptive agnosia) or affects access to semantic information about objects (associative agnosia).

Auditory agnosias are considerably rarer than visual agnosias for several reasons. Firstly the organisation of auditory system means that both sides of space are represented in both hemispheres, therefore selective damage to one hemisphere is unlikely to result in auditory recognition deficits as the other hemisphere can compensate. Secondly, damage to auditory regions is rare because of their location within the temporal lobes in relation to blood supply. Third, speech deficits may mask impaired non-verbal sound recognition, or are judged more important than non-verbal sound impairments (Polster & Rose, 1998).

Despite these factors, perceptual or semantic impairments for sounds do exist. Vignolo (1982) investigated semantic knowledge of sounds, using the meaningful sound recognition test, in which participants matched environmental sounds to their pictures. Perceptual knowledge was assessed using the meaningless sound recognition test, in which participants indicated whether two nonsense sounds were either the same or different. Vignolo (1982) found that patients with left hemisphere damage had impairments on the meaningful sound recognition test, implying that these patients had a deficit of access to semantic information about sounds, but no impairment on the meaningless sound recognition test. Patients with right hemisphere damage displayed the reverse relationship of impairments, suggesting that such patients have impaired perceptual processing of sounds. Though supported by some studies (Buchtel & Stewart, 1989; Schnider, Benson, Alexander, & Schniderklaus, 1994; Vignolo, 2003), other research has found little association between lesion side and type of auditory agnosia (Clarke, Bellmann, DeRibaupierre, & Assal, 1996; Tanaka, Nakano, & Obayashi, 2002).

Auditory agnosia often involves impairments of verbal recognition (i.e. pure word deafness), environmental sound recognition (i.e. non-verbal agnosia) or music recognition (i.e. amusia) and such impairments which can either co-occur with or present in isolation from each other (Peretz et al., 1994; Polster et al., 1998).

Zatorre and Halpern (1993) asked patients with left or right temporal lobe damage to either imagine or listen to familiar tunes and state whether a specific lyric was higher or lower in pitch than a reference lyric. Patients with right hemisphere damage had both perceptual and imagery impairments, whereas patients with left hemisphere damage had no impairments.

Although rare, some studies report similar dissociations in environmental sound perception abilities (e.g., impaired perception: Tanaka et al., 2002; preserved perception: Hattiangadi et al, 2005). Therefore these studies show that dissociations between verbal and non-verbal sounds may also exist after brain lesion.

Some studies report dissociations in deficits for living and non-living sounds, similar to the studies of neural activation to different sound categories. Kolinsky et al.(2002) found that their patient had impaired animal sound recognition, compared to object sounds. Other authors corroborate these findings (Hart & Gordon, 1992; Sirigu, Duhamel, & Poncet, 1991; Gainotti & Silveri, 1996).

This review has so far shown that though auditory imagery is largely a phenomenological experience, that imagery can affect perception and vice versa. Auditory imagery relies on similar areas of the brain as perception, can interfere with and be interfered with by auditory perception, and selective impairments according to sound categories can result from brain injury.

The review which follows focuses on schizophrenia and schizotypy, specifically in reference to hallucinations as auditory imagery has been implicated in the development of such experiences.

7. Schizophrenia

Kraeplin (1919) first described schizophrenia as a psychiatric disorder, which he called ‘dementia praecox’. He characterised this disorder as a chronic long-term condition of mental decline, starting in late adolescence. The symptoms included hallucinations and delusions, as well as a combination of cognitive, emotional and physiological symptoms. Bleuler (1911) coined the term ‘schizophrenia’ to reflect a division between thoughts and feelings. Bleuler noted that schizophrenic patients did not necessarily suffer the long-term decline that Kraeplin categorised, but that they can recover to some extent. Bleuler’s approach also noted four core features of this disorder, loosening of associations (difficulty organising thoughts), ambivalence in relating to others, autism (difficulty in relating to the milieu) and inappropriate affect. These symptoms were assumed to occur in all patients[1].

Later, Schneider's (1959) classification of first-rank symptoms for a diagnosis of schizophrenia reflected these symptoms, namely hallucinations, thought disorders and delusions. Schneiderian first-rank symptoms are still reflected in the current diagnosis system for schizophrenia. Diagnosis is now based on criteria set out in Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). DSM-IV specifies that a diagnosis is appropriate if at least two of the following symptoms have been experienced in the past month: hallucinations, delusions, though disorder (such as disorganised speech) and negative affect (such as blunted affect). In addition the patient should have experienced social or occupational disruptions, for at least 6 months. DSM-IV also features a number of subtypes of schizophrenia, dependant on the dominance of particular symptoms. Clinicians often use the Structured Clinical Interview to diagnosis whether a patient meets these criteria. This interview gathers as much detail as possible about the patient's current psychiatric symptoms, and features different modules applying to a range of disorders (e.g., psychotic, mood, substance abuse etc.). The psychotic disorders module features criteria for diagnosing schizophrenia and schizophrenia-like disorders. Two studies assessed the reliability of the SCID and confirmed it to be a robust method of diagnosing schizophrenia (Skre, Onstad, Torgersen, & Kringlen, 1991; Williams et al., 1992).

Studies also use these criteria to determine the prevalence of schizophrenia in the general population. The NEMESIS study found that overall 41.2% of their population

[1] See Boyle (1992) for a review of these studies

of 7076 people in the Netherlands had had a psychiatric diagnosis in their lifetime, with 0.4% meeting the diagnostic criteria for schizophrenia in their lifetime (Bijl, Ravelli, & van Zessen, 1998). Though earlier estimates suggested that the prevalence of schizophrenia is about 1% of the normal population (DSM-IV) other studies support that the prevalence is actually between 0.4 - 0.5% across different populations (Torrey, 1987; Goldner, Hsu, Waraich, & Somers, 2002; Saha, Chant, Welham, & McGrath, 2005).

As mentioned above there are a number of symptoms associated with schizophrenia, such as hallucinations, delusions and negative affect. A large focus of research has focused on the underlying construct of these symptoms in schizophrenia.

Crow (1980) proposed that schizophrenia involves two syndromes, which have separate pathology and are defined by the existence of positive or negative symptoms. Positive symptoms are in excess of normal mental function, and include hallucinations and delusions, which Crow (1980) associated with an increase in dopamine receptors. Negative symptoms are deficient compared to normal mental function, and include blunted affect and poverty of speech, which Crow (1980) associated with structural changes in the brain, particularly surrounding the ventricles (Andreasen, 1982a). Crow (1980) posited that despite the differing pathology, positive and negative symptoms can co-occur in patients. Later work by Andreasen (Andreasen, 1982b; Andreasen, 1984) supported this model and developed scales to measure positive and negative symptoms, (Scale for the Assessment of Positive Symptoms, SAPS and Scale for the Assessment of Negative Symptoms, SANS). The SANS and SAPS contains 50 items grouped into nine subscales (five in the SANS and four in the SAPS). From assessment of these symptoms Andreasen concluded that schizophrenia involves two different syndromes, such that patients can be classified as having positive, negative or mixed syndromes. They discovered a negative correlation between positive and negative symptoms, suggesting that such symptoms exist at either end of the same continuum.

The Positive and Negative Symptom Scale (PANSS, Kay, 1987) also assesses patients from the position of a dichotomous model of schizophrenia. PANSS is a 30-item structured clinical interview. Seven items focused on positive symptoms, seven focused on negative symptoms, whilst the remaining 16 items measured general psychopathology. PANSS assessment similarly supports the notion of a two dimensional model, as again a negative correlation was found between positive and negative scales.

Later studies have however suggested that there may be more dimensions to schizophrenia than just positive or negative syndromes. For instance Liddle (1987) assessed 40 patients with a DSM-III diagnosis of schizophrenia, with the SAPS and SANS. Factor analysis revealed three factors: a delusional-hallucination syndrome, a disorganised syndrome and a negative syndrome. Other factor analyses supported a similar three-factor model (Arndt, Alliger, & Andreasen, 1991; Peralta, Deleon, & Cuesta, 1992) rather than the positive–negative dichotomy suggested by Crow (1980) and Andreasen (1982). Other studies however suggest an even more complex factor structure, positing further dimensions and sub-dimensions, of the basic three factor model (see Peralta & Cuesta, 2001 for a review of such studies).

Despite the considerable debate regarding the key symptom structure of schizophrenia, many studies focus on determining the cause of these symptoms in schizophrenia.

For instance as mentioned above, Crow (1980) suggested a biological component to the development of schizophrenia, with increased dopamine receptors associated with positive symptoms and neurological impairments associated with negative symptoms.

Frith & Done (1988) model also implicated neurological differences in the development of schizophrenia symptoms. Their theory posited that negative symptoms arise through impaired connection between the prefrontal cortex and basal ganglia, leading to difficulty in action initiation (and therefore poverty of speech and disorganised speech). In contrast this theory hypothesised that impaired connections between the prefrontal cortex and hippocampus result in difficulties in discriminating internally or externally generated events (and therefore leading to the manifestation of hallucinations and delusions).

In contrast Hemsley's (1993) neuropsychological model suggested that schizophrenia symptoms originate from a difficulty in separating memories from sensory input, as a result of impairments within the hippocampus and surrounding areas. This leads to automatic activation of irrelevant memories, which (due to their irrelevant and unintended nature) are concluded to be alien and from an external source, leading to hallucinatory experiences. Also hippocampal impairments lead to problems distinguishing relevant and irrelevant details in the environment. As such, attention is drawn to irrelevant details, leading patients to determine why their attention is drawn to the item, resulting in delusional beliefs of meaningfulness.

Many studies suggest a genetic component to schizophrenia development. Gottesman and Sheilds (1982) estimated that the risk of developing schizophrenia increases 10-fold for siblings of patients with schizophrenia, and there is about a 50% of a person developing schizophrenia for monozygotic twins of patients, or if both parents have the disorder. In addition Kety, Rosenthal, Wender, & Schulsin (1971) found higher rates of schizophrenia-related illness in families of adopted schizophrenics than control participants. Rosenthal, Wender, Kety, Welner, & Schulsin (1971) found that 31.6% of adopted-away children from schizophrenia-spectrum diagnosis patients developed schizophrenia compared to 17.8% for the control group.

These models therefore suggest that schizophrenia symptoms (particularly positive symptoms) arise through biological or neurological differences of schizophrenic patients. Such models do not explain why some people with biological predeterminants of schizophrenia do not develop the disorder. Other models posit that combination of biological and other factors are involved in the development of schizophrenia.

For instance the Stress-Vulnerability model (Zubin & Spring, 1977) posited that biological differences create an underlying vulnerability to schizophrenia, which under stress conditions, results in manifestation of psychosis symptoms. Patients vary in the extent of their vulnerability, so that lesser or greater amount of stressors result in disorders varying in severity. This explains the variability of schizophrenia between patients. Evidence for this theory comes from studies that find that patients that develop psychosis symptoms have an increase in stressful life events in the months prior to onset (see Norman & Malla, 1993for a review of such studies) for a review of such studies). This model does not explain the expression and variability of specific symptoms however.

Indeed, attempts to define a theory that encompasses all symptoms of schizophrenia have generally been unsuccessful (Birchwood & Jackson, 2001), therefore many theoretic approaches focus on particular symptoms (e.g., delusions, hallucinations) or symptom clusters (e.g. negative symptoms).

In particular, a number of cognitive models have been designed to describe the development of positive symptoms. For instance Maher (1974) posited that delusions develop in association with hallucination, such that they are reasonable explanations for anomalous hallucinatory experiences. Further models posit more complex mechanisms however.

Garety, Kuipers, Fowler, Freeman, & Bebbington (2001) view delusions as a failure of reasoning processes that enable people to consider alternate hypotheses for unusual thoughts and experiences. As such deluded patients tend to jump to conclusions, leading them to be more willing to accept their initial hypotheses for unusual events without seeking further confirmatory evidence.

Bentall, Kinderman, & Kaney (1994) posited that delusions develop through an exaggeration of normal cognitive biases. For instance Kaney & Bentall (1989) administered the Attributional Styles Questionnaire to delusional patients, depressed patients and controls. Delusional and depressed patients both attributed to global and stable consequences to negative events (i.e. will affect a wide range of events, and will be present in the future). The two groups differed in the perceived internality of events however, such that depressed patients viewed negative events as internally-caused (e.g. due to the self) and positive events as externally-caused (e.g. due to luck), whereas deluded patients displayed the opposite beliefs. Deluded patients resembled normal participants in their beliefs of externality/internality for good and bad events, however in an exaggerated form. Bentall et al. (1994) posited that if attributional biases reflect attempts to protect self esteem in normal personality, delusions indicate an extreme attempt to protect self esteem in deluded patients.

Similar cognitive theories have been developed for hallucinations also (see Hallucination section below). Therefore it appears that a symptom-based approach to understanding schizophrenia is more explanatory than a single unified construct. As mentioned above however, there is also some disagreement about the factor structure of schizophrenia symptoms, namely what are the key symptom clusters. This provokes the question of whether schizophrenia as a unified concept is valid at all. A number of researchers have suggested that rather than taking a syndrome-based categorical approach to assessing schizophrenia, a dimensional approach is more appropriate for assessment, based on individual symptoms (Bentall, 1992; Bentall, 2006; Allardyce, Gaebel, Zielasek, & van Os, 2007).

Bentall (1992) questioned the concept of schizophrenia stating that it should meet a number constructs to be considered reliable and valid. Firstly measurement of symptoms using different criteria should result in concordant diagnoses for schizophrenia to be a reliable construct. In contrast, Brockington, Kendell, and Leff (1978) found low concordance rates in diagnosis when comparing different criteria. In addition for schizophrenia to be a valid construct, it should be associated with a specific set of symptoms (e.g., delusions, hallucinations, negative affect etc). It is rare however, to find two ‘schizophrenic’ patients who present with the same symptoms. Also the major symptoms of schizophrenia are associated with other psychiatric disorders. Specifically, schizophrenia, bipolar disorder and depression have similar characteristic, such as affective symptoms, mania or delusions. Therefore some studies focus on the degree of association between these disorders, to corroborate whether they are actually distinct. Everritt (1972) performed cluster analysis on 480 patients diagnosed with a psychiatric disorder. This analysis clusters participants together based on the similarities between their symptoms. Therefore all patients with schizophrenia should have formed a distinct cluster from those with bipolar disorder for instance. Though the analysis found separate clusters for these disorders, about 60% of the patients fell into one or more poorly defined clusters that do not apply to a specific disorder, questioning the dissociation between different syndromes.

In addition for schizophrenia to be a valid construct it should be discontinuous with the normal personality (Allardyce et al., 2007) such that one either has schizophrenia or not. However as shown in the next section, people within the normal population do present with psychosis-like symptoms, suggesting a continuity between ‘normality’ and schizophrenia.

8. Schizotypy

The early Kraeplin concept of schizophrenia suggested a clear boundary between mental illness and mental wellness, such that one can either be schizophrenic or not. Later theories posited that experience of schizophrenic symptoms may exist on a continuum, with normal mental health at one end and full blown schizophrenia at the other (Strauss, 1969; Crow, 1995).

Associated with this theory is the concept of ‘schizotypy’, in which individuals can have similar traits to that of schizophrenia, but at a subclinical level (Meehl, 1962). Two models describe the occurrence of schizotypy (Claridge, 1997). The quasi-dimensional model (a predominately medical perspective) views schizotypy as a mild form of actual schizophrenia, whereas the fully dimensional model (predominately psychological perspective) views schizotypy as an extreme variation of normal personality. A key difference between these theories is that the fully dimensional model suggests that it is possible to be a healthy schizotype, whereas the quasi-dimensional model sees schizotypy symptoms as underlying schizophrenia itself.

Meehl (1962), a proponent of the quasi-dimensional model posited that many people have an inherited neural defect, which in partnership with environmental influences, results in behaviour characteristics similar to schizophrenia. These include ‘cognitive slippage’ (a mild form of thought disorder) interpersonal problems and anhedonia. Meehl (1962) termed this condition ‘schizotaxia’ and people with such characteristics ‘schizotypic’. This theory supported a Stress-Vulnerability viewpoint and posited that many people have latent schizophrenic symptoms, which only develop into full-blown schizophrenic symptoms with exposure psychosocial stressors. Claridge (1992) likened this situation to hypertension, where blood pressure ranges along a continuum. This only causes problems for the individual when it becomes very high, and exacerbated by stress to cause a stroke.

Indeed, onset and relapse of psychosis is associated with stressful life events, such as daily life stressors (Myin-Germeys, Delespaul, & van Os, 2005) child abuse (Read, van Os, Morrison, & Ross, 2005) and trauma (Morrison, Frame, & Larkin, 2003).

Other studies focussed on the inherited aspect of schizotypy, investigating the existence of such symptoms in relatives of schizophrenic patients. Such studies find that first-degree relatives of schizophrenic patients have more schizotypal traits than control participants, such as disorganised thinking, negative symptoms or social dysfunction, participants (Kendler, Mcguire, Gruenberg, & Walsh, 1995; Kendler, Thacker, & Walsh, 1996).

High risk participants (i.e. those with a genetic risk of developing schizophrenia through having a relative with the disorder) may show early-signs of psychosis development in childhood. For instance studies of high risk children show that neurological abnormalities (Keshavan et al., 1997; Sharma et al., 1997), minor physical abnormalities (Schiffman et al., 2002) and language abnormalities (Bearden et al., 2000) were associated with schizophrenia development in adulthood. In addition, psychosis-like symptoms in childhood are predictive of future psychiatric diagnosis. For instance Poulton et al. (2000) found an association between that reports of psychotic experiences at 11 years old, such as voice hearing and delusional thinking and schizophreniform diagnosis 26 years later. Also Escher et al. (2002)

found that 40 % of 12 year old children who heard voices, still heard them 3-years on, with negative voice appraisals and anxiety/depression symptoms distinguishing those who still heard the voices from those who did not. Therefore such early signs of psychosis symptoms are predictive of future psychosis development, suggesting a continuity of symptoms from childhood to adulthood.

In response to Kety et al. (1971) and Rosenthal et al's. (1971) study findings (see above) Spitzer et al. (1979) designed Axis-II of DSM-III to incorporate diagnosis of "subclinical schizophrenia". Axis II disorders focus on personality disorders, including schizotypal personality disorder (SPD). This disorder is characterised by a combination of social difficulties and unusual experiences, but not at levels that impair occupational functioning or involving chronic hallucinations or delusions (these can occur at a less severe level). DSM-III features 9 criteria for diagnosis of SPD, including magical thinking, social anxiety and odd behaviour. In addition clinicians use a specific SCID for personality disorders to diagnosis SPD (SCID-II, Spitzer, Endicott, & Gibbon, 1979).

Fully dimensional models posit that schizotypy is an extreme variation of normal personality. This theory orginates from Eysenck's theory of psychoticism (1960), who suggested that people vary along a continuum of psychoticism, and can remain healthy, despite having traits of psychoticism.

Questionnaire-based self-report measures are often used to assess schizotypy from this perspective. The focus of such measures varies between questionnaires, with some incorporating all symptoms (i.e. O-LIFE, Mason, Claridge, & Jackson, 1995;

SPQ, Raine, 1991) and others on more specific symptoms, such as the Perceptual Aberrations Scale (Chapman, Chapman, & Raulin, 1978) or the Launay-Slade Hallucination Scale, (Launay & Slade, 1981).

Similarly to schizophrenia, much debate has focused on the factor structure of schizotypy, and used factor analysis to determine its structure. The predominant view posits that two factors underlie the nine SPD symptoms, namely cognitive-perceptual dysfunction (e.g., magical thinking, unusual perceptions, etc) and deficient social functioning (e.g., social anxiety, no close friends etc) (Siever & Gunderson, 1983) . Other studies posit more than two factors however. For instance Bentall et al. (1989) conducted a large scale investigation of 14 schizotypy scales. These scales and the Eysenck personality questionnaire formed the 'Combined Schizotypal Traits Questionnaire' and factor analysis was performed on the data. Exclusion of a delusion scale revealed a three-factor model relating to positive symptomatology, negative symptomatology and the third relating to social anxiety and cognitive disorganisation. With the delusion scale included a four-factor model emerged containing the same three factors as the previous model, plus an extra factor of asocial behaviour. A later study confirmed this factor structure (Claridge et al., 1996).

The Schizotypal Personality Questionnaire (SPQ, Raine, 1991) focuses on all nine of the DSM-III criteria of SPD. It was developed from other schizotypy questionnaires, including the SANS (Andreasen, 1984), and SCID-II (Spitzer et al., 1979). Factor analysis of the SPQ found a three-factor model of schizotypy, containing a cognitive-perceptual factor (reflecting unusual perceptual experiences and magical thinking), an intrapersonal factor (reflecting social anxiety and lack of close relationships) and a disorganisation factor (reflecting disordered speech or behaviour) (Raine et al., 1994). Fossati et al. (2003) confirmed this factor structure in an Italian sample, and established that this structure was consistent across age and gender.

Later, Mason, Claridge and Jackson (1995) developed the Oxford-Liverpool Inventory of Feelings and Experiences through exploratory factor analysis of several questionnaires and scales measuring personality traits, schizotypy and psychosis symptoms, such as the Launay-Slade Hallucination Scale (Launay et al., 1981), the Psychotism scale from the EPQ (Eysenck & Eysenck, 1975), the MMPI (Golden & Meehl, 1979) and the STQ (Claridge & Broks, 1984). The final 104-item questionnaire contained four scales. The Unusual Experiences (UE) scale assessed positive symptoms, such as hallucinations and delusions. The Introvertive Anhedonia (IA) scale questioned about negative symptoms, such as flattened affect and lack of social involvement. The Cognitive Disorganisation (CD) scale assessed attention, concentration and decision-making difficulties. These subscales reflect the three factor structure of schizophrenia symptoms (Mason & Claridge, 2006), however the O-LIFE also contained a fourth subscale. The Impulsive Nonconformity (IN) subscale assessed self-abusive, violent or reckless behaviours. Analysis of the final questionnaire revealed good internal reliability and high inter-correlations between the subscales (only the correlation between UE and IA was not significant) (Mason et al., 1995; Mason et al., 2006). In addition each subscale correlated with age, showing that scores generally decrease with age. Factor analysis revealed a four factor model, relating to the original four scales of the questionnaire.

Therefore factor analyses revealed that a similar factor structure underlies schizotypy as schizophrenia, supporting the theory that they exist on a continuum with more extreme symptoms. A problem with continuum approaches is that the separation between schizotypy and schizophrenia becomes based on an arbitrary cut-off based on a group average score. In borderline cases, it becomes difficult to establish whether a set of symptoms constitute a mental illness or an eccentric personality. It is difficult to treat schizotypy and schizophrenia based on a continuum approach, as such intervention is based on a dichotomous decision (i.e. to treat or not to treat). Decisions of this sort may lead to stigma associated with a diagnosis, and revert mental illness to something one either has or has not.

The next section of this review focuses on theories of hallucination development as well as measurement of hallucinations.

9. Hallucinations

Investigation of hallucination is an area where imagery research may be most pertinent, since such experiences involve miscategorisation of internal mental events (i.e. imagery) as external.

DSM-IV defines hallucinations as:

"A sensory perception that has the compelling sense of reality of a true perception but that occurs without external stimulation of the relevant sensory organ."

Hallucinations are most commonly associated with psychosis, but are a common symptom of other psychiatric disorders such as PTSD, bipolar disorder and depression

(Beck & Rector, 2003). Indeed Brasic (1998) noted over 40 other disorders that involve hallucination.

Hallucinations can be experienced in all modalities (e.g., auditory, visual, tactile etc) but auditory hallucinations are by far the most commonly experienced by schizophrenic patients, followed by visual, tactile and olfactory/gustatory hallucinations (Mueser, Bellack, & Brady, 1990). Schneider (1959) noted auditory hallucinations as first-rank symptoms of schizophrenia. Indeed estimates range between 50 % and 74% for patients lifetime prevalence of auditory hallucination (Slade & Bentall, 1988; Baethge et al., 2005).

Auditory hallucination vary greatly in their presentation, ranging from simple sounds such as tapping or scratching noises, to full sentences from a number of different voices (Beck et al., 2003). Schneider reported three main types of hallucination: Hearing own thoughts out loud, hearing voices arguing and hearing voices commenting on one's actions. In addition the features of hallucinated voices vary greatly. Nayani and David (1996) investigated the phenomenology of auditory hallucination. Hallucinations varied in their perceived location with the majority of patients stating that they originate from the head, and a smaller percentage experiencing them as from an external source. The majority of voices were also male and often known to the patient in real life.

Hallucinatory experiences of patients are often associated with negative symptoms. For instance Krabbendam et al. (2005) found that depressed patients have a greater incidence of hallucinations when followed-up 3 years later, than non-depressed patients. In addition periods of stress or trauma often mark onset of hallucinations (Beck et al., 2003).

Though hallucinations are a common symptom of psychiatric disorders, many people who do not suffer from such disorders report experience of hallucinations. Romme and Escher (1989) appeared on a television show, in which they invited people who heard voices to participate in a study of voice hearing. A large number of people who responded to the invitation did indeed experience voices, but about 40% claimed they coped well with them, considered them a normal part of their life and did not need treatment for them.

Because reports of non-psychiatric hallucinatory experiences are so common, a number of studies focused on ascertaining the proportion of the normal population that have such experiences. Estimates ranged from 5% and 25% of people in the normal population who experience hallucinations (Tien, 1991; Barrett & Etheridge, 1992; Young, Bentall, Slade, & Dewey, 1986), with the average around 10% (Slade et al., 1988). Indeed Slade and Bentall (1988) suggest that hallucinations exist at the extreme end of a continuum, with normal vivid mental imagery being at the other end.

10 Measurement of hallucinations

Many measures investigate hallucination experience as part of wider assessment of general positive and negative symptoms of schizophrenia or schizotypy (e.g., PANSS, STQ and O-LIFE) whereas others focus purely on hallucination.

The PSYRATS (Haddock, McCarron, Tarrier, & Faragher, 1999) focuses on positive symptoms, containing both a hallucination scale and a delusion scale. The scales investigate the duration, frequency and amount of distress associated with hallucinations and delusions, which are rated by the clinician. Investigation of ratings of 71 patients revealed good inter-rater reliability, and significant correlations between some of the PSYRATS items and another psychiatric assessment scale, suggesting that the scale was both reliable and valid. Drake et al. (2007) confirmed this in a later study, finding it was beneficial for assessment of first-episode psychosis patients. This scale was developed for assessment of psychiatric patients however, and as such cannot be used to assess hallucination-like experiences in non-psychiatric patients.

In contrast the Launay-Slade Hallucination Scale (Launay et al., 1981) is a self-report measure, which can be used with psychiatric and non-psychiatric patients. Originally tested on prison inmates, this scale determined whether the incidence of hallucination was greater in these participants than in the normal population. The final scale featured 12-items, covering experiences ranging from vivid mental events to hearing unusual sounds in the absence of an external source, to which participants gave yes or no responses. As well as confirming a higher incidence of hallucination-like experiences in the inmate group, factor analysis of the whole data-set revealed two main factors: “tendency to hallucinatory experiences” and “negative response set” (only two items loaded significantly on the latter factor, as they were negative in tone). In addition scores on the LSHS correlated with that of Eysencks Psychotism scale (P) suggesting that this scale had good reliability.

A later study by Bentall and Slade (1985) investigated the reliability of the LSHS, by asking 150 male non-psychiatric participants completed the LSHS on two separate occasions. This study used the same scale as in the previous study except that participants rated the extent to which they agreed with each item. They found a correlation between the two versions of the questionnaire, suggesting that the LSHS measure is reliable and measures a stable trait.

Other LSHS studies focussed on the factor structure of this questionnaire in non-psychiatric participants. Aleman et al (2001) conducted a Principal Components Analysis (PCA) on the LSHS scores of normal participants and found three main factors: “general hallucination tendency”, “subjective externality of thought” and “vividness of daydreams”. Another PCA with psychiatric patients revealed similar factors of “vivid daydreams”, “clinical auditory hallucinations” and “intrusive thoughts”, and also an extra factor of “subclinical auditory hallucinations” (Levitan, Ward, Catts, & Hemsley, 1996).

The above studies suggest that the LSHS is a reliable and valid method of measuring hallucinatory experiences, in both the psychiatric and non-psychiatric populations. This scale does not take into account unusual sensory experiences in modalities other than the auditory modality, however. It also does not explicitly measure factors related to the severity of hallucination experiences such as distress and distraction caused by such experiences, or the frequency at which the experiences occur. Such measures would help determine how pervasive hallucination-like experiences are.

The Cardiff Anomalous Perceptions Scale (CAPS, Bell, Halligan, & Ellis, 2006) is a 32-item scale which questions about anomalous experiences in all sensory modalities, without direct questions about schizotypy or hallucinatory experiences. It featured nine categories of questions related to different anomalous experience. Participants give yes or no responses to the questions and where participants gave a positive response, they also rated the experience using a 5-point rating scale, on three dimensions: the distress, the intrusiveness, the frequency of the experience. Bell et al. (2007) tested this scale with a group of non-psychiatric participants and a psychiatric group of participants.

As well as good internal reliability, the CAPS total score correlated well with the O-LIFE, LSHS and PDI, suggesting that this questionnaire has good construct validity.
A PCA on the CAPS total score revealed three factors: “temporal lobe experiences”, “chemosention” (i.e. unusual olfactory and gustatory experiences), and “clinical psychosis”. Therefore it appears that the CAPS encompassed a wide range of experiences which may influence hallucination disposition.

These measures are useful tools for determining the presence and pervasiveness of hallucinations. The following section turns to theories of the development of hallucinations.

10. Theories of Hallucination

Research into the development of hallucinations mainly suggests that internal representations of sounds or voices are confused with perceptions for actual sounds or voices. Theories differ in the mechanism by which this occurs. Three major cognitive theories of hallucination development are reviewed here. These are imagery theory, inner speech theory and reality monitoring theories.

10.1 Imagery vividness theories of hallucination

Mintz and Alpert (1972) argued that patients who suffer from hallucinations have abnormally vivid imagery, so, relative to other individuals images of patients are more ‘perceptual-like’ in quality and therefore more easily confused for actual percepts.

Mintz and Alpert’s (1972) grouped hallucinating and non-hallucinating schizophrenic patients according to whether they scored high or low in an auditory imagery vividness task. A greater number of hallucinating patients scored highly on this task compared to non-hallucinating patients. Participants also took part in an auditory detection task in which they reported the presence of sentences in white noise and their confidence in their responses. Relative to non-hallucinating participants, hallucinating participants had a lower correlation between their accuracy in sentence report and their confidence ratings. Mintz and Alpert (1972) argued that the hallucinating group were relatively poor at judging their own internal state, leading them to be overly confident in the task. One theory is that ‘reality testing’ or ‘source monitoring’ deficits are a major mechanism in developing hallucinations, in combination with high imagery vividness. Mintz and Alpert (1972) suggested that hallucinations come about when participants have abnormally vivid imagery plus a deficit in the mechanism which determines these events as internal, leading to confusion between mental images and external events.

Sack et al. (2005) also investigated the association between vividness of imagery and hallucinations in patients with psychosis. Hallucinating patients had high imagery vividness in all modalities, compared to non-hallucinating patients. There was no relation between severity of symptoms and imagery vividness however, suggesting that vivid imagery may be an independent trait of patients but not causally related to hallucination.

Increased imagery vividness of patients who hallucinate is not a consistent finding, however. Brett and Starker (1977) and Starker and Jolin (1982) both found that hallucinating participants reported less vivid imagery than non-hallucinating and control participants. These authors have suggested that patients who hallucinate have abnormally low vividness of imagery which leads them to assume that any experience of imagery is something extraordinary (Smith, 1992).

Studies using imagery ratings may be prone to effects of social desirability (on the one hand) or lack of motivation or apathy (on the other) however. To overcome such problems, later studies have investigated the association between hallucination and more objective measures of auditory imagery. Bocker, Hijman, Kahn and de Haan (2000) had groups of hallucinating, non-hallucinating and control participants complete tests of visual and auditory perception, objective measures of vividness of imagery and reality discrimination tasks. Hallucinating and non-hallucinating participants did not differ in performance on the perceptual tests though there were differences between hallucinating and non-hallucinating participants on a visual imagery-perception interaction task. Hallucinating participants benefited more from imagining a visual target, when asked to detect a faint visual stimulus in visual noise, compared to non-hallucinating participants. Few differences were apparent between hallucinating and non-hallucinating participants in the reality discrimination task, though a hallucination severity and response bias in an auditory sound detection task correlated. In other studies though (Aleman, Bocker, Hijman, de Haan, & Kahn, 2003) hallucinating and non-hallucinating patients do not differ on a number of objective imagery tasks.

More recent studies have investigated the links between imagery and perception in participants who have schizotypic traits and/or hallucination-like experiences, but at a subclinical level. Research with such participants is advantageous because it removes the confounding affect of medication or cognitive deterioration, allowing investigation of characteristics of schizophrenia without these confounds. There is a suggestion that normal mental imagery lies on a continuum with more perceptual like experiences, such as hallucinations, with vivid imagery lying somewhere between the two extremes (Slade et al., 1988). Merckelbach and van de Ven's (2001) studied provided evidence for this, as their normal participants reported hallucination-like experiences in an auditory listening task. They told participants the song 'White Christmas' may be present in a sample of white noise. Despite the absence of the song in the white noise throughout the experiment, 32% of participants believed that they heard the song. Therefore experience of "hearing" imaginary sounds actually appears to occur frequently in the normal population.

Barrett (1993) compared imagery vividness across seven modalities for groups of high or low hallucinatory participants from the normal population . High hallucinators had more vivid imagery on several subscales suggesting a common underlying actor contributing to imagery vividness; however there was no association between vividness of auditory imagery and occurrence of auditory hallucinations.

Aleman et al. (1999) investigated both objective (imagery triad comparison) and subjective measures of imagery (QMI) in high and low hallucination prone participants. Similarly to previous studies with hallucinating patients, high hallucination prone participants had more vivid imagery on the QMI compared to low hallucination prone participants, whereas the opposite relationship occurred for the objective imagery task. A later study again compared objective and subjective imagery vividness. This experiment included more objective tasks however, such as a visual letter imagery task, an auditory music imagery task and an imagery-perception interaction task, and the imagery triad task (Aleman, Nieuwenstein, Bocker, & de Haan, 2000). Hallucination proneness did affect visual imagery vividness ratings but the only effects on the objective tasks emerged through higher correlations (and smaller overall differences) between the imagery and perceptual versions of the triad task for the high hallucination prone participants. The study also revealed a significant correlation between imagery and perception on the music task for the high hallucination prone participants. These data suggest that high hallucination prone individuals may have a greater similarity between imagery and perception.

Overall these studies show that the pattern of performance of high hallucination prone, normal participants is similar to that of hallucinating psychosis patients. However it appears that though high hallucination prone and hallucinating participants are more likely to rate their imagery as more vivid, such vividness does not necessarily translate to better performance on objective measures of imagery vividness.

10.2 Inner Speech Theories of Hallucination

Inner speech is the normal process of internal verbalisation of our thoughts. The inner speech theory suggests that hallucinations originate from a disrupted inner speech generation, such that the patient does not realise that these thoughts are self-generated and so categorises them as ‘alien’ (Frith and Done, 1987).

Allen et al (2006) provided evidence for this theory, by using a source monitoring task. In this task, high and low hallucination prone participants distinguished their own distorted voices from other voices, showing that high hallucination prone participants made more errors about the source of the voices than low hallucination prone participants.

Related to the theory is the suggestion of increased subvocalisation use in hallucinating patients. Subvocalisation often accompanies inner speech and involves covert use of vocal organs, as if preparing the person for actual speech. Gould (1949) amplified a patients subvocalisations and found that the resulting speech matched closely with what the patient claimed the voices had been saying. Other studies confirm that that hallucinating patients experience subvocalisation during hallucinations (Mcguigan, 1966), and that interfering subvocalisation can interrupt hallucination (Green & Kinsbourne, 1990).

Neuroimaging research also supports the inner speech theory. McGuire et al. (1996b) and Shergill et al (2000) found that inner speech is associated with activity in the left inferior frontal lobe, and that patients have increased activity in this region during hallucinations.

Other studies however, found no impairment in the inner speech processes of hallucinating patients. According to Baddeley's (1986) working memory model, inner speech utilises the phonological loop to maintain information in a short-term verbal store. If hallucinations do involve a disruption to this system, then patients who hallucinate should be impaired at performing concurrent tasks that also use the phonological loop. This was the premise of Haddock (1996) study, who asked hallucinating patients to perform an immediate serial recall task using phonologically similar items to the targets, and compared their performance to non-hallucinating and control participants. Though the psychiatric group did perform poorer than controls, hallucinating and non-hallucinating patients did not differ from each other, suggesting that hallucinations are not the product of disrupted phonological loop function.

10.3 Reality monitoring theories of hallucination

Source monitoring is the process by which attributions of the source of a memory are made. Studies into this investigated how people determine old events from new. Many researchers hypothesise that source monitoring processes are disrupted in hallucinating patients.

A number of paradigms assess source monitoring abilities, including word association and signal detection tasks. In the word association task, participants generate associate words for the target items. The participant is then presented with their own associated words, and some experimenter-generated associate words, to which they were asked to discriminate the self-generated items. These studies show that hallucinating and high hallucination prone normal participants are more likely to attribute their self-generated items as experimenter generated, indicating a bias towards externalising internally created information (Morrison & Haddock, 1997; Brebion et al., 2000).

Reality monitoring is a subcategory of source monitoring, which is also hypothesised to be disrupted in hallucinating patients. This process enables discrimination of real from imaginary events (Johnson & Raye, 1981). In normal people found that these processes vary as a function of a variety of cognitive factors. For instance Johnson et al (1977) ran a paired associates learning task, in which participants gave the associated pair item when given a cue item. The authors varied the number of times each pair was presented and tested, and at the end of the experiment participants estimated how often they were tested on each item. The number of times each item was presented influenced participant's judgements, suggesting that the process of internalising the information (i.e. studying the items) altered the participants ability accurately judge how often they were tested on an item. Further studies have determined that reality monitoring is also influenced by the familiarity of the stimuli, vividness of the memory for the item and the relevance of the stimuli (Johnson et al., 1981; Johnson, 1997; Johnson, Taylor, & Raye, 1977). Other studies focused on the role of reality monitoring processes in the development of hallucinations (Bentall & Slade, 1985; Rankin & Ocarroll, 1995).

Sound detection tasks are often used to assess reality monitoring abilities. In such tasks, participants detect faintly presented sounds in white noise, and the minimum threshold for detection, as well as number of false positive responses are analysed. A large number of false positives indicate an externalising bias, suggesting difficulty in distinguishing internal from external events.

For instance Rankin and O'Carroll (1995) grouped normal participants according to whether where they scored on a hallucination proneness questionnaire, and asked them to detect sounds in noise. The high hallucination prone group had a greater bias for believing in the presence of the sounds, but the two hallucination proneness groups did not differ from each other in sensitivity. In addition, participants completed a paired-associates task in which participants imagined the associated word for each test item. During this task, the experimenter varied the number of times each item was presented and at the end of the test, the participant estimated how often each item appeared. High and low hallucination prone participants did not differ in their frequency estimations for the items they actually heard, but high hallucination prone participants had higher frequency estimation for imagined items than low hallucination prone participants. Therefore these tasks show that those prone to hallucination-like experiences differ from those who do not in their metacognitive abilities, such as reality monitoring, rather than differing in underlying sensory processing.

Bentall and Slade (1985) compared hallucinating and non-hallucination patients on sound detection of word presented in white noise. Hallucinating patients has a stronger bias to responding positively than non-hallucinating participants. There was however no difference between the two groups in their sensitivity to the target. The findings were similar for comparisons between normal (i.e. non-psychiatric participants) high and low hallucination prone participants. Bentall and Slade suggested these findings were due to a reality testing deficit, rather than due to differences in imagery vividness between the two groups.

A similar study determined the neural underpinnings of reality monitoring deficits (Barkus, Stirling, Hopkins, Mckie, & Lewis, 2007).This study compared high and low hallucination prone participants ability to detect a word in white noise. In addition a subset of high hallucination prone participants did this during fMRI scanning. The behavioural results were very similar to Bentall and Slade's; there was an increased response bias for the high hallucination prone participants but no difference between high and low hallucination prone in target sensitivity. The fMRI findings revealed activation in the right middle temporal gyrus, bilateral fusiform gyrus and the right putamen, during false positive responses. Barkus et al. (2007) noted that previous studies report activity in these regions during investigations of auditory hallucination, mental imagery, and auditory detection tasks for speech stimuli in the past. They also found activation in the right frontal areas, bilateral temporal regions and the left cingulate gyrus, when false positives were contrasted with hit responses. They interpreted frontal activations as related to speech and task difficulty, and the temporal lobe activations to auditory hallucination experience and speech perception. Finally they suggested that the cingulate activity was a result of auditory processing, and a link between cerebellum activity and making task decisions in ambiguous circumstances (Barkus et al., 2007). These patterns of activation suggest an association between false positives in normal participants and actual hallucination experiences, and may also involve mental imagery.

11 Emotion, schizophrenia and auditory hallucination

Negative symptoms are another key feature of schizophrenia, which can affect both emotional experience and recognition. Several studies demonstrated impairments in schizophrenic patients at judging emotion from faces (Edwards, Jackson, & Pattison, 2002; Morrison, Bellack, & Mueser, 1988) and voices (Edwards et al., 2002). Blunted affect and depression are common in schizophrenia and a there is a suggested a link between negative emotional state and cognitive-perceptual disturbances, such as delusions and hallucinations (Serper & Berenbaum, 2008).

Many studies focus on the association between emotional content and schizophrenia symptoms. For instance Haddock found that patients with thought disorder symptoms, exhibited more symptoms when interviewed with emotional salient questions (i.e. “what do you worry about?”), than emotional neutral (“tell me about a famous sportsman?”). Emotional content also interacts with delusions. Bentall and Kaney (1989) used an emotional Stroop task, in which the response latency to name the colours of threat-related, depression-related and neutral words. Deluded patients took significantly longer to perform the task for threat-related words, suggesting that they grabbed the attention of such patients more. These studies suggest that emotional content can exacerbate schizophrenia symptoms.

Freeman and Garety (2003) also reported an association between experience of hallucination and emotional disturbances, such as depression and anxiety. As well the association with negative mood, hallucinations are often emotional in content. Nayani and David (1996) conducted a phenomological survey of hallucinations, and found that 60% of the patients in their sample had abusive hallucinations. There is some variations in the content of hallucinations however, with some patients finding them distressing and frightening, while others find them pleasant and amusing. Chadwick and Birchwood (1994) and Birchwood and Chadwick (1997) posited that the level of distress evoked by hallucinations may be the result of relationship the patient has with their voices, rather than their frequency or content. This model suggests that beliefs about the voices are predictive of the distress that they cause. Patients who believe that their voices are powerful and uncontrollable, experience more distress regardless of how often they hear voices or whether they were malevolent or benevolent.

Morrison and Haddock (2002) investigated the effect of emotional content on source monitoring abilities of hallucinating, non-hallucinating and normal control participants. Participants received a list of positive, negative and neutral words and to each they generated a related word. After a delay, participants received the original and self-generated words again and asked whether each word was from the original list, or from the self-generated words. Hallucinating patients were more likely to attribute emotional words as generated by the experimenter, suggesting greater bias of hallucinating participants to externalise emotional material

Laroi, Van der Linden and Marczewski (2004) supported this finding, using a procedure similar to Morrison et al. (2002). High and low hallucination prone participants did not differ in their accuracy at judging whether items were from the old or new list, but high hallucination prone participants were much more likely to indicate that the items were experimenter- rather than self-generated. In addition such errors were greatest for negative words, followed by positive and neutral words. Laroi et al. suggested the emotional words increased arousal more in the high hallucination prone participant, leading to disturbance of source monitoring.

Brain imaging research suggests a link between externalisation in hallucination-prone patients and controls, and increased neural activation in response to emotional information. Sanjuan et al. (2007) asked patients with and without hallucinations to listen to emotional and neutral words doing fMRI scanning. Hallucinating patients showed stronger activation in the left middle superior temporal lobes, orbitofrontal cortex, temporal cortex, insula, cingulate, and amygdale when hearing emotional compared to neutral words. Previous studies revealed activation of these regions when investigating hallucination and emotional processing. Activation strength did not differ between emotional and non-emotional words for controls participants. The data indicate that hallucinating patients may experience greater arousal when hearing emotional words, which may contribute to the externalising bias, associated with emotional memories.

12 Thesis Outline

The studies of imagery vividness reviewed above suggest that auditory imagery vividness is not purely a subjective experience or a measure of social desirability, but can interact with and affect perception (Allbutt et al., 2008; Baddeley & Andrade, 2000; Cui et al., 2007; Farah & Smith, 1983; Segal et al., 1970; Tinti, Cornoldi, & Marschark, 1997). Chapter 2 extends previous findings about imagery vividness and the interaction between imagery and perception. Experiment 1 investigated the influence of different cognitive factors on imagery vividness ratings, by manipulating sound category, familiarity and imagery cues. Experiment 2 investigated the effect of sound perception on imagery vividness ratings for different sound categories. This determined whether perception of a sound can influence auditory imagery vividness of another sound. Experiment 3a further examined the relationship between imagery vividness and sound familiarity to determine whether the two were independent of each other. Finally Experiment 3b assessed the influence of cognitive factors such as vividness, familiarity and detection cues on ability to detect faintly presented sounds in white noise. This experiment therefore aimed to determine the extent to which imagery can influence perception.

Chapter 3 investigated the association between imagery and perception for non-verbal sounds in normal participants, using fMRI. This study focused on animal and environmental sounds, since there is little previous research using these categories. Assessment of the neural response to animal and environmental sounds also naturally examines the brain structures associated with semantic differences between sound categories, as animal and environmental sounds are more similar to each other in acoustic characteristics, compared to other sound categories (see Chapter 3 for further discussion). The current study built upon previous studies that have found differences in neural activation in response to animal and environmental sound categories, and extended it to investigate the overlap between imagery and perception for such sounds. Experiment 4 assessed imagery and perception for animal and environmental sounds, using a sparse-sampling fMRI design, which allowed for imagery to occur during silent gaps in the scanning sequence. From this, the association in activation between imagery and perception was assessed, along with (i) the association between activation to animal and environmental sounds and (ii) the association between self-rated imagery vividness and neural activation, without the confound of scanner noise.

Chapter 4 had two aims. Firstly it compared the sample of participants in the current thesis to that of previous studies, to establish similarity in terms of schizotypy traits. Secondly it assessed the relationship between imagery vividness and measures of schizotypy. The studies reviewed above suggest an association between auditory imagery vividness and hallucinatory experiences. This is a contentious issue as some studies do find links between subjective measures of imagery and hallucination proneness and others do not. Such studies have varied in their assessment of imagery, hallucination proneness and participants assessed. Therefore this chapter aimed to clarify the relationship between these measures.

Chapter 5 expanded on the findings in Chapter 2, by investigating how the interaction between imagery and perception differs in participants who vary in hallucination proneness. Studies of the interaction between imagery and perception suggest that imaging a target affects the detection of sounds in noise (Farah et al., 1983; Segal et al., 1970). In addition studies suggest that the response bias to responding that sounds are present in noise is greater in participants who are prone to hallucinations (Barkus et al., 2007; Bentall et al., 1985). Experiment 6 employed the same sound detection task as used in Chapter 2. This experiment assessed the effect of hallucination proneness on sound detection, to determine firstly how this variable influenced sensitivity and response bias to the target sound. Secondly the study determined how sound vividness and familiarity interacted with hallucination proneness to affect detection. Finally it determined how the availability of a detection cue interacted with hallucination proneness to influence sound detection.

Chapter 6 also built upon Chapter 2 by further examining how different auditory stimuli can influence sound detection of participants who vary in hallucination proneness. From the studies reviewed here it appears that emotional processing impairments are a key facet of psychosis, and emotional content can interact with hallucination to exacerbate impairments in other processes, such as source monitoring. Few studies have investigated the effect that hallucination proneness and/or source monitoring abilities have on either imagery vividness or the detection of auditory stimuli with an emotional connation however. Experiment 7a used auditorily presented emotional and neutral words to investigate the affect that emotional connotation had on imagery vividness and memory of high and low hallucination prone participants. Experiment 7b investigated whether high and low hallucination prone participant differ in the influence of emotional content on sound detection.

In conclusion the first part of this thesis investigates some of the mediators of auditory imagery vividness, and how vividness affects the interaction between imagery and perception. It also assesses the little researched categories of animal sounds and environmental sounds, to discover whether there is a semantic dissociation between these sound categories. In addition an investigation of the neural correlates of vividness ratings determined whether higher imagery ratings correlate with activation changes within the brain. The thesis also examines the relationship between hallucination proneness and auditory imagery vividness, to determine if there is an association between these factors. Sound detection tasks using meaningful sounds and words investigated the interaction between imagery and perception in high and low hallucination prone participants, and to determine how different stimuli types affect this interaction. These studies clarify the nature of auditory imagery vividness by showing that it can be influenced by a number of cognitive and personality characteristics.

Chapter Two. Vividness of auditory imagery and sound detection

Auditory imagery vividness is a measure of subjective imagery strength, which requires participants to rate how clearly they can imagine sounds. This chapter determined the effect that different factors have on vividness, and how cognitive factors influence detection of sounds in noise. Experiment 1 investigated the effect of sound category, familiarity and cues to imagine sounds (i.e. pictures vs. cues) on imagery vividness ratings. Experiment 2 also assessed the effect of presenting an auditory judgment task during the imagery encoding period, on vividness ratings. Both experiments revealed differences in vividness dependant on sound category and familiarity level. In Experiment 1 picture cues resulted in higher imagery ratings than name cues and in Experiment 2 listening to a sound compared to white noise resulted in lower vividness ratings. Experiment 3a obtained familiarity ratings for each sound, as well as test-retest reliability for vividness ratings. Vividness ratings were reliable and vividness and familiarity positively correlated, though using familiarity rating as a covariate did not remove the effect of sound category on vividness ratings. Finally Experiment 3b investigated the theory that vivid images have a more ‘perceptual-like’ quality, using a sound detection task. The study investigated sensitivity, response bias and confidence in hearing familiar and unfamiliar high vividness (i.e. music and speech) and low vividness sounds (i.e. animal and environmental sounds) in noise. Valid cues or no cue was given, as to the identity of the target in noise. Vividness and familiarity affected the bias, but not on sensitivity to the target. Detection cues affected confidence ratings only. The current study therefore demonstrated that cognitive factors can affect auditory imagery vividness and in turn, these factors can interact with and affect detection of sounds in noise.

Introduction

Studies investigating auditory imagery have found similarities between imagery and perception in sensitivity to the acoustic characteristics of stimuli (Intonspeterson, 1980; Intonspeterson, Russell, & Dressel, 1992) and in their neural activations for different categories of sound (Bunzeck, Wuestenberg, Lutz, Heinze, & Jancke, 2005; Halpern & Zatorre, 1999; Halpern, Zatorre, Bouffard, & Johnson, 2004; McGuire et al., 1996a; Shergill et al., 2001).

Other studies have found that imagining a sound can have an inhibitory (Segal et al., 1970) or facilitatory (Farah et al., 1983) effect on the detection of a faintly presented tone in noise (see Chapter 1 for full description of these studies). Further, Segal and Fusella (1970) found that imaging an unfamiliar sound impaired tone detection more than a familiar sound, suggesting that greater processing capacity was required to generate the unfamiliar image, leaving less available to detect the tone.

Baddeley and Andrade (2000) found that imagining nonsense speech was associated with lower imagery vividness ratings than imagining meaningful speech. Also Kraemer et al. (2005) found a greater association between silent gaps in familiar music and stronger activation in the left auditory association areas than silent gaps in unfamiliar songs, suggesting generation of auditory imagery during these gaps for familiar but not for unfamiliar sounds. These studies suggest that the familiarity of a sound can affect the strength of the auditory imagery and sound detection.

No studies have looked at sound familiarity and imagery abilities for other sound categories, however. The aim of this chapter is to investigate how sound category and familiarity affects auditory imagery vividness and detection of sounds in noise.

The present study used both ratings of auditory imagery and measures of detection under imagery conditions to assess how auditory imagery varies across different stimuli and as a function of stimulus familiarity. If vividness ratings reflect the perceptual experience of imagery, then there should be convergence across the different measures.

Experiment 1 investigated how cues to imagine sounds affect auditory imagery for different types of familiar and unfamiliar sound. The study examined auditory imagery vividness in response to a name cue compared to a picture cue. Lehmann and Murray (2005) found that forming a multisensory association (i.e. auditory-visual) at encoding improved memory for target items. It is possible that cross modal vision-sound associations also form part of our representations of stimuli (see Humphreys & Forde, 2001, for one account). These associations may be formed more strongly for pictures and sounds than for words and sounds, given that the visual occurrence of the object is more likely to co-occur with the sound of the object than the name for the object. It follows that auditory imagery might be invoked more strongly by a picture than a word cue.

Experiment 2 employed a dual task procedure to investigate how auditory imagery can be affected by performing another auditory task during the encoding period. Baddeley and Andrade (2000) found that articulatory suppression during encoding resulted in lower imagery vividness ratings for music, compared to performing a visuo-spatial suppression task during encoding. Also Tinti et al. (1997) found that sound detection had a selective disrupting effect on auditory imagery whereas visual detection had little effect on sound memory. These data suggest that processing in auditory imagery and perception rely on a common, limited capacity resource, distinct from the resources required for visual perception. We have little information, though, about whether the resource holds across all types of sound, or whether separate pools of resource are recruited for different sounds. Imaging evidence for separate neural loci for different sounds raises the possibility that there are distinct resources however. Here we assessed whether there were selective interference effects for different sources of sound.

The measure used in Experiments 1 and 2 was a five-point imagery vividness scale, which required participants to rate the vividness of their auditory image from '1' being "no image at all" to '5' being "perfectly clear and vivid as normal hearing". This measure assesses individual differences in imagery experience as it asks people to introspectively consider their imagery abilities. These measures are used frequently in visual imagery research and in questionnaires about imagery abilities.

As mentioned in the literature review however, opinion is mixed with regards to the validity of vividness ratings. Some studies find good correlation between vividness and objective measures of memory (Cui et al., 2007; Marks, 1973), whereas others report no correlation (Chara & Hamm, 1989). Previous studies measuring auditory imagery vividness often aim to determine differences between imagery abilities of schizophrenic and non-schizophrenic participants (Aleman et al., 2000; Aleman, Bocker, & de Haan, 2001). Thus there is a lack of understanding of factors affecting the vividness of auditory imagery in normal participants. Experiment 3a determined the reliability of vividness ratings and how they relate to sound familiarity. Participants made vividness ratings on two separate occasions and rated the familiarity of the sounds. This determined the validity of vividness ratings, their consistency across time and that they were not a function of the familiarity of the sounds, rather due to generation of auditory imagery.

Experiment 3b employed a sound detection task to investigate directly whether the vividness and familiarity of the image affected sensitivity and/or any bias towards responding positively (making false positives). As mentioned above mental imagery can either inhibit (Segal et al., 1970) or facilitate (Farah et al., 1983) sound detection, depending on the experimental conditions. The current experiment determined which of these theories is true, by manipulating both ability to imagine the target (i.e. through provision of an imagery cue) and by manipulating the vividness associated with the target (i.e. by providing high and low vividness sound for detection). The study hypothesised differences in detectability of the sounds, depending on the vividness and sound familiarity associated with the sound and on cue provision.

Experiment 1: Effects of cue type on auditory imagery vividness

Method

Participants

12 people from the School of Psychology at the University of Birmingham volunteered to participate in this experiment in return for credit. The participants' were all female and their mean age was 19.75. All stated that they had near perfect hearing in both ears.

Design

The task required participants to rate their auditory imagery for different sounds, upon presentation of a picture or written name cue

Stimuli

The stimuli were 96 2 – 3.5sec sound clips, with a sampling rate of 22500 Hz and a bit rate of 16 bits. These sounds were equally distributed across four categories of sound (animal sounds; environmental sounds; music; speech). Twelve sounds in each category were unfamiliar. Familiar animal and environmental sounds were collected from the Internet and edited using a sound-editing program (Audacity 1.2.6). Unfamiliar sounds were created by changing the pitch, speed, tempo or by reversing various aspects of the familiar versions. Familiar tunes were nursery rhymes, popular television and film themes, Christmas carols and well-known classical pieces, whereas unfamiliar versions were made-up tunes, played on a keyboard, clarinet, organ or recorder. The familiar and unfamiliar tunes did not differ significantly in the number of notes (*t(22) = 1.242, p = 0.387*). Speech clips were single words (each spoken twice) spoken by two men and two women. Familiar words were single syllable English words and unfamiliar words were single syllable non-words.

Ten participants, who did not take part in the following experiment, rated the familiarity of each of these sounds.

These familiarity ratings were analysed using a repeated measures ANOVA, with the factors being sound category and familiarity. A significant main effect of sound category, $F(3, 27) = 6.288$, $p = 0.015$, partial $\eta 2 = 0.411$, and Bonferroni adjusted post hoc analyses showed a trend for differences between animal sounds and environmental sounds, and animal sounds and music only, ($p=0.051$ and $p=0.074$ respectively). Also familiar items were indeed rated as being more familiar than unfamiliar items, $F(1,9) =169.375$, $p <0.001$, partial $\eta 2 =0.950$. The interaction between sound and familiarity was not significant, $F(3, 27,) = 1.521$, $p = 0.245$, partial $\eta 2 = 0.145$

The picture and name associated with each sound were shown during the presentation phase of each trial. Picture items for unfamiliar animal sounds were found on the internet (www.humandescent.com). They were photographs of ‘chimeric animals’ edited using ‘Photoshop’. Made-up names were associated with these animals, but these names did not differ in length from the familiar animal names. Unfamiliar environmental sounds were also created using ‘Photoshop’, and consisted of photographs of two different objects joined together. Again made-up names were associated with these items.

For the familiar music clips, the picture item was of the appropriate instrument and the name item was the name of the song in the music clip. The picture item for the unfamiliar sound clips was also a picture of the appropriate instrument, but the name items were made-up song titles. The pictures associated with the familiar and unfamiliar speech sound clips were photographs of the person speaking, and the name item was the word that was spoken.

Procedure

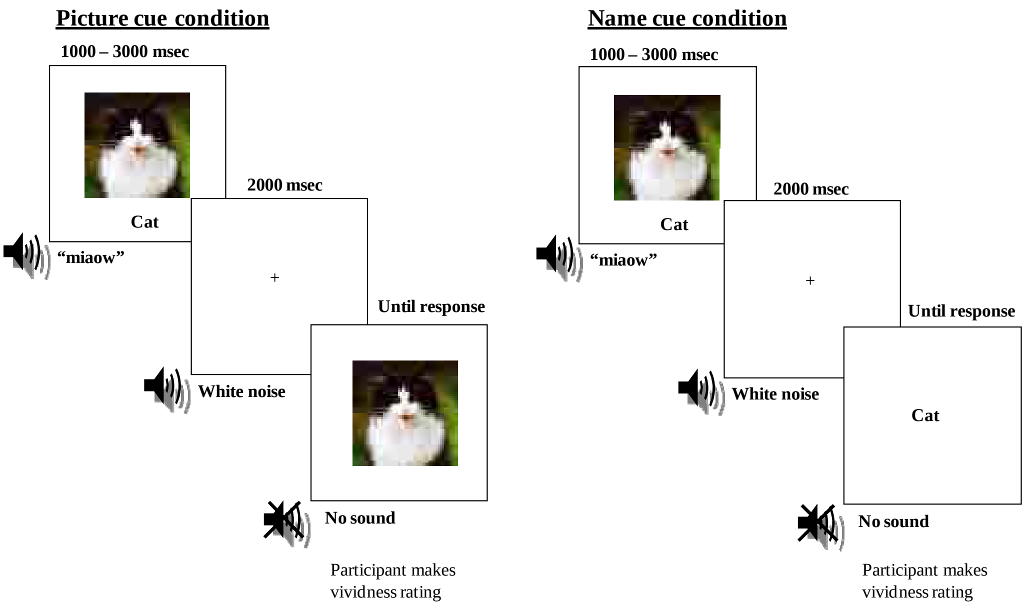


Figure 1: Experiment 1. Trial procedure for each condition

The experimenter informed participants that the task involved use of auditory imagery, and that they would hear familiar and unfamiliar sounds and make an imagery vividness rating for each, upon presentation of a cue (see Appendix A. for instructions given to participants). Figure 1 shows the trial procedure for this experiment.

The experiment was presented using E-prime (Psychology Software Tools, Pittsburgh, PA). On each trial, participants first heard a sound clip (e.g. a meowing cat) through a standard set of over-ears headphones and saw the picture and name (e.g. a picture of a cat, with "Cat" written underneath) associated with the sound on the computer screen. They then heard 3.5sec of white noise to clear the sensory memory for the sound item, to ensure image generation for the item upon presentation of the imagery cue.

Participants then saw the picture or the name of the item on the screen and rated the strength of their imagery for the sound of the item by pressing a number from one to five, using the keyboard (see

Table **1**). Participants heard each sound clip twice throughout the experiment, once with a picture cue, and once with a name cue. All conditions were randomised. Participants completed a practice session of four trials, and then if they were happy with the task, they started the full experiment which lasted approximately 40 minutes.

Table 1: Imagery vividness rating scale

Rating	Statement
1	No image at all, you only "know" that you are thinking of the object
2	Image is vague and dim
3	Image is moderately clear and vivid
4	Image is clear and reasonably vivid
5	Image is perfectly clear and as vivid as normal hearing

Results

Analysis focused on two aspects of the imagery task: imagery vividness ratings and the reaction times for rating imagery vividness.

Imagery Vividness Ratings

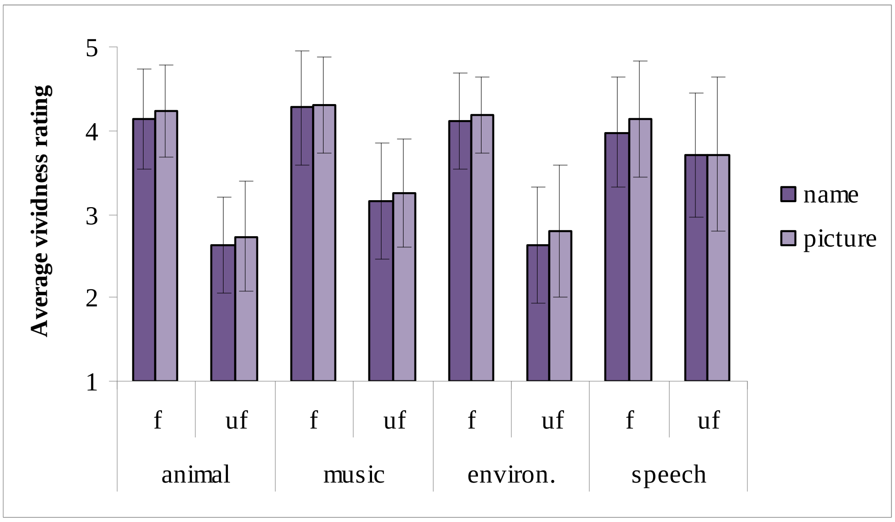


Figure 2: Average imagery vividness rating for different familiar (f) and unfamiliar (uf) sounds when cued with a name or a picture.

Figure 2 shows the mean imagery vividness ratings for each sound in each condition. Imagery vividness ratings were analysed using a 4 x 2 x 2 repeated measure ANOVA. The factors were sound category (animal sound, environmental sound, music, speech), familiarity (familiar vs. unfamiliar) and cue type (name vs. picture).

The main effect of sound category was reliable, $F(3,33) = 6.789$, $p = 0.005$, *partial* $\eta^2 = 0.382$: vividness was highest for music imagery, followed by speech imagery, animal sound imagery, with the lowest ratings given to environmental sound imagery. Bonferroni adjusted post hoc analyses revealed a significant difference between environmental sound imagery and speech only ($p = 0.042$) however. Familiar sounds received higher ratings than unfamiliar sounds, $F(1, 11) = 86.177$, $p <0.001$, *partial*

$\eta^2 = 0.887$. Picture cues received higher ratings than name cues, $F(1, 11) = 5.164$, $p=0.044$, *partial* $\eta^2=0.319$.

The interaction between sound category and familiarity was significant, $F(3,33) =29.533$, $p<0.001$, *partial* $\eta^2=0.729$. For familiar sounds there was no significant difference between the sound categories, $F(3, 33) = 0.938$, $p = 0.386$, *partial* $\eta^2 = 0.079$. For unfamiliar sounds however there was an effect of sound category, $F(3,33) =1.859$, $p<0.001$, *partial* $\eta^2=0.665$. Bonferroni adjusted post hoc analyses revealed lower ratings for animal sounds compared to music *(p=0.045)* and speech *(p=0.001)* and for environmental sounds compared to music *(p=0.043)* and speech *(p<0.001)*. Music items received lower ratings than speech items (*p=0.012)*. No other interactions were significant.

Reaction Times (RTs) to make imagery vividness ratings

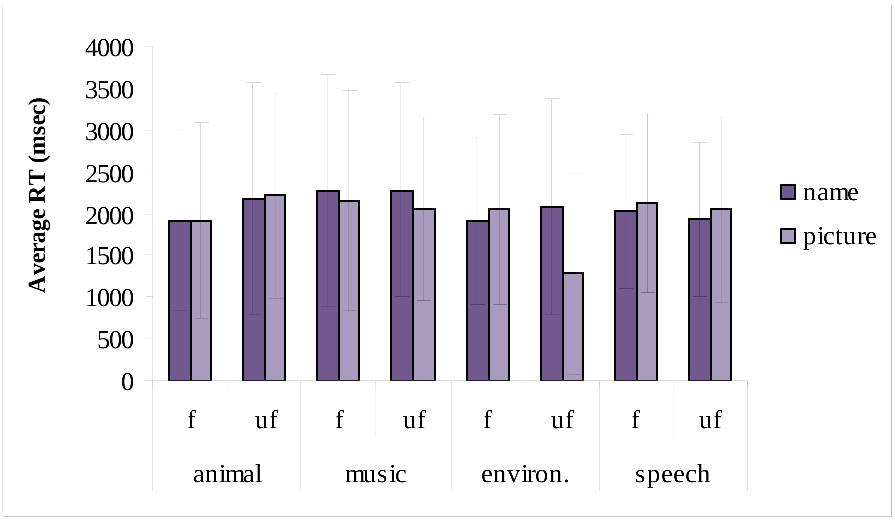


Figure 3: Average RTs (msec) to rate the familiar and unfamiliar sounds when cued with a name or a picture

Figure 3 shows the average time that participants took to make their imagery vividness ratings in each condition. These RT's to make the imagery vividness ratings were analysed using a 4 x 2 x 2 repeated measure ANOVA, with the following factors: sound category (animal sound, environmental sound, music, speech); familiarity (familiar vs. unfamiliar); cue type (name vs. picture). This revealed no significant main effects or interactions.

Discussion

This experiment investigated imagery vividness ratings for different categories of familiar and unfamiliar sounds, following either a name or picture cue. Analysis of familiar sounds revealed that there was little difference between sound categories. For unfamiliar sounds however, animal and environmental sounds received lower ratings than music and speech sounds. This may simply be because we have more experience of unfamiliar tunes or words, therefore they are less unusual than unfamiliar animal and environmental sounds. This may certainly be so for music, as unfamiliar music received higher familiarity ratings than other unfamiliar sounds. However unfamiliar speech did not differ from either animal sounds or environmental sounds in rated familiarity, suggesting that increased experience of certain categories of unfamiliar sounds cannot be the only contributing factor to differences in imagery vividness between sounds.

Participants gave higher vividness ratings when imagining the sounds in response to a picture cue compared to a name cue. This is probably due to forming a stronger association between visual and auditory items, as they are likely to appear together. This offers support for Lehmann and Murray’s (2005) finding that visual memory was better for items encoded as visual-auditory memories than for those encoded as verbal-auditory memories. The current study therefore suggests that not only memory but also auditory imagery vividness is stronger with visual-auditory associations, showing that memory and imagery processes may behave in similar ways.

When questioned, many participants stated that they were able to rehearse the image of the sound, despite presentation of white noise to disrupt this process. Therefore it is difficult to judge, whether any category differences reflect the initial evocation of the sound or the ability to maintain it. Experiment 2 examined whether introducing a masking noise varying in similarity to the imagined sound selectively disrupted maintenance of auditory images.

Experiment 2: Auditory interference effects on auditory imagery vividness

The second experiment focused on the effect of distracter sounds on imagery vividness. This experiment used a dual task procedure, so that following presentation of the target sound, participants made a familiarity judgement about another sound, and then made their imagery rating for the target. The secondary task aimed to prevent rehearsal of the stimulus, so that image generation would only occur when cued. The task required participants to indicate whether the intervening sound was familiar or unfamiliar, to ensure that participants paid attention to it. The sounds presented in the familiarity task were either congruent to the target sound (i.e. from the same sound category) or incongruent to it (i.e. from a different sound category). Would interference be category specific?

Items in the same sound category share semantic and some acoustic properties (e.g., pitch, tempo, etc.) therefore we expected that imagery vividness ratings would be lower when a congruent sound followed the target compared to an incongruent sound, as rehearsal of the target sound competes for resources with the interference sound. We expected unfamiliar sounds to receive lower imagery vividness ratings, in accordance with the previous experiment.

Method

Participants

18 participants volunteered to do this study, recruited from the School of Psychology at the University of Birmingham. The participants' ages varied from 18 to 21 years, and the group was predominantly female (15 female and 3 male). All rated themselves as having near perfect hearing in both ears.

Design

As in Experiment 1, the imagery task required participants to rate their auditory imagery for different familiar and unfamiliar sounds. 24 familiar and 24 unfamiliar target sounds from each of the four sound categories were presented (16 per interference condition: congruent, incongruent and white noise). Animal and environmental sounds and pictures were collected and edited in the same way as before. Familiar and unfamiliar speech clips were single English and non-words spoken twice, by two male and two female native English speakers, and the pictures were of people speaking. Familiar and unfamiliar music clips were familiar and created tunes played on the piano, clarinet, organ or recorder, and the pictures were of the appropriate instrument.

Participants heard each target sound once in the experiment, so that prior presentation would not affect subsequent imagery ratings. The type of interfering sound following each sound category was counterbalanced across participants, so participants heard all sounds the same number of times with each interference type. Sixteen familiar and 16 unfamiliar sounds were presented from each sound category as the congruent and incongruent interference sounds (again each presented once). In the control condition, 3.5sec of white noise was presented during the maintenance period.

A separate experiment assessed the familiarity of each target sound. Ten independent participants rated their familiarity with each sound, and ratings were analysed with a 4 x 2 repeated measures ANOVA (factors: sound category and familiarity). Sound category was a significant main effect, $F (3, 27) = 10.683, p = 0.003, \text{partial } \eta 2 = 0.543$. Bonferoni post hoc analyses revealed that familiarity was lower for animal sounds than music, environmental sounds and speech $(p<0.05)$; the other sound categories did not differ from each other. Ratings for familiar sounds were higher than for unfamiliar sounds, $F (1, 9) = 167.142, p <0.001, \text{partial } \eta 2 = 0.949$. There was also an interaction between sound and familiarity, $F (3, 27) =5.235, p = 0.015, \text{partial } \eta 2 = 0.368$. This showed a significant difference between sound categories when the sound was familiar, $F (3, 27) =20.087, p < 0.001, \text{partial } \eta 2 = 0.691$, and a trend towards a significant difference for unfamiliar sounds, $F (3, 27) = 3.616, p = 0.063, \text{partial } \eta 2 = 0.287$. Bonferoni post hoc analyses of familiar sounds revealed lower familiarity for animal sounds than music, environmental sound and speech $(p < 0.001)$.

Imagery Questionnaire

In addition, and extra to Experiment 1, each participant in the main experiment completed an imagery questionnaire, which aimed to accustom participants to the task of self-evaluating their auditory imagery, and to obtain a measure of long-term memory for sound.

The imagery questionnaire had a visual imagery section and an auditory imagery section. Each section consisted of 24 items. The visual imagery section contained items from the 'Vividness of Visual Imagery Questionnaire' (Marks, 1973). This questionnaire requires participants to form an increasingly complex visual image of four items and rate the vividness of their image, using a five-point scale (one being the lowest rating, i.e. "no image" and five being the highest rating, i.e. "image is perfectly clear, like normal hearing"). An auditory imagery section was designed because previous imagery questionnaires either do not have a dedicated auditory imagery section (VVIQ, Marks, 1973) or are outdated (e.g. shortened version of QMI, Sheehan, 1967). This section was designed in a similar way to the VVIQ, so that participants formed increasingly more complex auditory images of four items (Sheehan, 1967). The items in this section focused on the four sound categories (animal sounds, environmental sounds, music and speech). Participants imagined the sound and rated each item for auditory imagery vividness using the same five-point scale as the visual imagery section. In addition eight single simple items were included at the beginning of each questionnaire, to ease participants into the task of rating their imagery vividness (see Appendix E for the Auditory Imagery Questionnaire).

A pilot investigation of 27 participants determined the reliability and validity of this questionnaire. Each participant completed the questionnaire twice, once in each session, separated by a week. Overall, the visual imagery section of the questionnaire had a Cronnebach Alpha of 0.887 and the auditory imagery section had an overall Cronnebach Alpha of 0.885, suggesting that both sections had good internal reliability.

The questionnaire scores from each session correlated well for both the visual imagery and auditory imagery sections, *(*$r(27)=0.595$, $p<0.001$ and, $r(27)= 0.707$, $p<0.001$ respectively*)*, showing good test-retest reliability.

In addition there was a good correlation between the visual imagery subsection and the auditory imagery subsection, $r(27)=0.713$, $p<0.001$. Previous research has found a good association between these factors, suggesting a general imagery vividness factor (Allbutt et al., 2008). Therefore the good association between these factors here suggests that the auditory imagery section (i.e. the new subsection) has good construct validity.

In the following experiment, half the participants completed the imagery questionnaire before the experimental trials and half afterwards.

Procedure

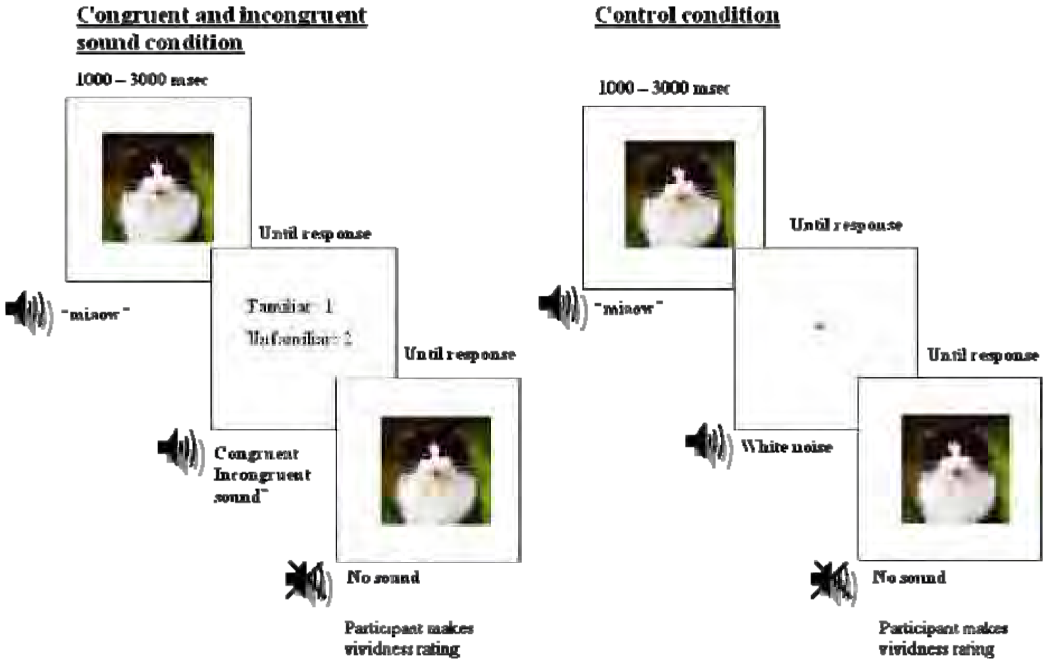


Figure 4: Experiment 2. Trial procedure in each condition.

Figure 4 shows the trial procedure for each condition of this experiment. The experimenter informed participants that the task involved rating their auditory imagery vividness for different familiar and unfamiliar sounds, following a sound familiarity judgement task (see Appendix B for instructions given to participants). Each trial involved presentation of picture and sound of the target item, and following a maintenance period, participants rated the vividness of their imagery for the target sound. During the maintenance period participants made a familiarity judgement in response to either a congruent sound (i.e. same category as the target) or an incongruent sound (i.e. different category as the target). The control condition involved presentation of white noise during the maintenance period. Participants completed the control condition either before or after the interference conditions. Each participant had four practice trials to ensure that they understood the task, and then they completed the full experiment, which lasted approximately an hour.

Results

Imagery Vividness Ratings

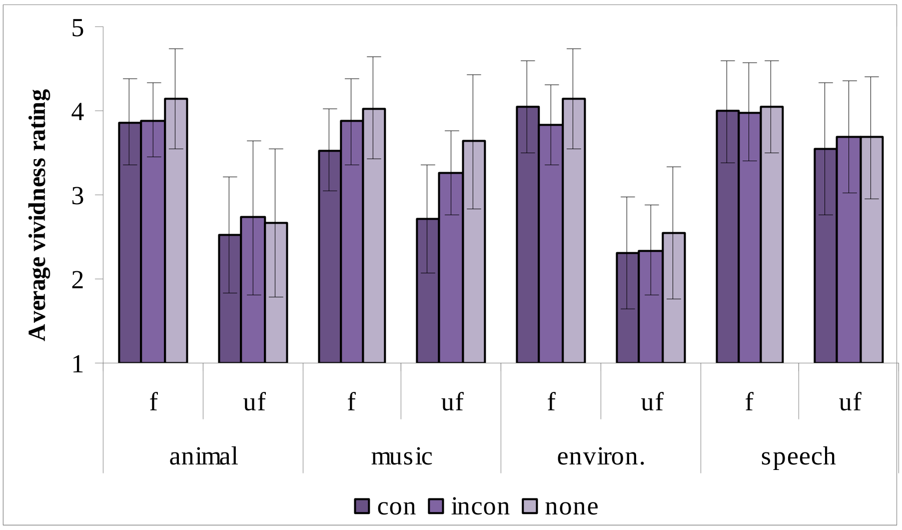


Figure 5: Average imagery vividness ratings for different familiar and unfamiliar sounds, with different types of interference

Figure 5 shows the average imagery vividness rating for the congruent, incongruent and control conditions. Analysis of imagery vividness ratings consisted of a 4 x 2 x 3 repeated measures ANOVA. The factors were sound category (animal sound, environmental sound, music and speech) familiarity (familiar vs. unfamiliar) and interference type (congruent, incongruent or white noise). There was a significant main effect of sound category, $F(3, 51) = 12.135, p<0.001, partial\ \eta^2=0.417$. Bonferroni adjusted post hoc analyses revealed significant differences between speech and animal sounds ($p=0.004$), environmental sounds ($p<0.001$) and a borderline difference between speech and music ($p=0.053$). Familiar sounds received higher vividness ratings than unfamiliar sounds, $F(1, 17) = 147.026, p<0.001, partial\ \eta^2=0.896$. The main effect of interference was also reliable, $F(2, 34) = 5.905, p = 0.017, partial\ \eta^2=0.258$. Bonferroni adjusted post hoc analyses revealed significant differences between congruent and incongruent interference ($p<0.05$) and congruent and white noise interference (*p*<0.05) but not between incongruent interference and white noise (*p*>0.05).

There was also a significant interaction between sound category and interference, $F(6, 102) = 8.471, p<0.001$, *partial* $\eta^2=0.333$. When analysed further there was a reliable contrast between the different interference conditions for music only, $F(2, 34) = 16.695, p<0.001$, *partial* $\eta 2=0.495$. A Bonferroni adjusted post hoc analysis showed significant differences between congruent interference and incongruent interference only ($p<0.001$) for this sound category.

The interaction between sound category and familiarity was also significant, $F(3, 48) = 28.106, p<0.001$, *partial* $\eta^2=0.637$. Sound categories did not differ significantly when familiar, $F(3, 51) = 1.514$, $p = 235$, *partial* $\eta 2 = 0.082$, but unfamiliar sounds, $F(3, 51) = 25.081, p<0.001$, *partial* $\eta 2 = 0.596$. Again ratings were lower for animal and environmental sounds, compared to music and speech. Bonferroni adjusted post hoc analysis revealed significant differences between animal sounds and environmental sounds ($p=0.043$), animal sounds and speech ($p<0.001$), music and environmental sounds ($p=0.002$), music and speech ($p=0.037$) and environmental sounds and speech ($p<0.001$). The pattern of these data replicates Experiment 1.

RT for making imagery vividness ratings

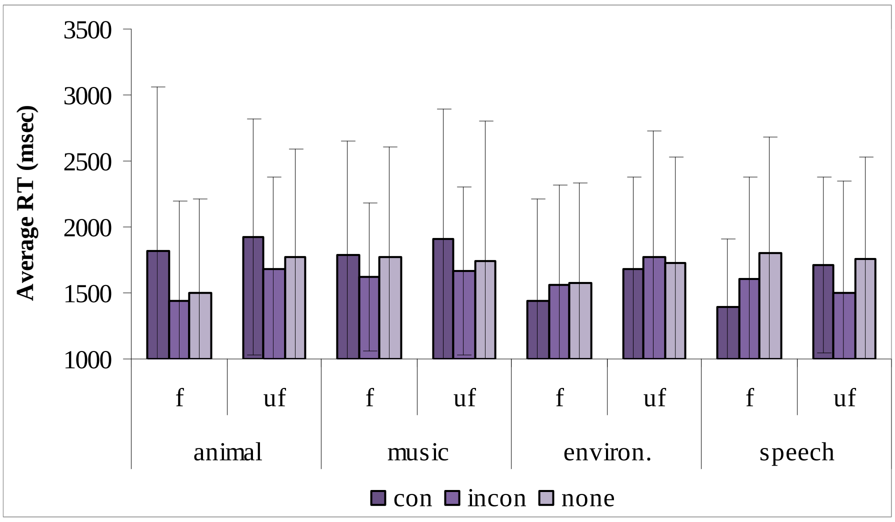


Figure 6: Average RT for different familiar and unfamiliar sounds, with different types of interference

Figure 6 shows the average time for participants to rate their imagery vividness for the congruent, incongruent and control conditions. Analysis of the RT data involved a 4 x 2 x 3 repeated measures ANOVA. The factors were sound category (animal sound, music, environmental sound and speech), familiarity (familiar vs. unfamiliar) and interference type (congruent, incongruent and white noise). There was a significant main effect of familiarity, $F(1, 17)=12.013$, $p=0.003$, *partial* $\eta^2=0.414$, and a significant interaction between sound category and familiarity, $F(3, 51)=3.645$, $p=0.032$, *partial* $\eta^2=0.161$. There was no reliable effect of sound category in analysis of each separate familiarity level. Finally there was a significant interaction between sound category and interference, $F(6, 102) = 2.645$, $p = 0.049$, *partial* $\eta 2 = 0.135$. Further analysis revealed a borderline significant difference between sound categories when the interference sound was congruent, $F(3, 51) = 2.822$, $p = 0.073$, *partial* $\eta 2 = 0.142$. It took longer to rate animal sounds and music compared to environmental sounds and speech with congruent interference, though Bonferroni adjusted post hoc analyses revealed no reliable differences.

Correlations

Correlations with imagery questionnaire

The correlation between the auditory imagery questionnaire score and the visual imagery questionnaire score was significant, $r\ (18) = 0.529, p = 0.024$. Participants who rated themselves as having high auditory imagery vividness also rated themselves as having high visual imagery vividness. The correlations between average imagery ratings in the ratings experiment and auditory imagery questionnaire scores were not significant however.

Correlation between sound length and interference type

In the experiment the length of each sound clip varied between 1.5 and 3.5sec. This was to ensure the sounds were as natural as possible, which meant that animal and environmental sounds often lasted a maximum of 2sec, while music sound clips often lasted a maximum of 3.5sec. To analyse whether this had an effect on the ratings for the items, the sound length for each item and the mean vividness rating for each item were correlated. This revealed no significant association between vividness and sound length, $r(128) = 0.128, p=0.077$.

In addition, during the familiarity judgement conditions, the time between when the participants heard the target sound and made their imagery ratings was not controlled. This was because participants gave their ratings immediately after their familiarity judgement. Therefore the average RT for making the familiarity judgment for each sound in the two interference conditions was also calculated. A paired-samples t-test revealed no significant difference in RT between the two conditions, $t(19) = 1.398, p$

= 0.164. Also there was no significant correlation between the average familiarity judgement RT and the average rating for each sound, *r (192) = -0.121, p=0.093*.

Similarity ratings

Assessment of the similarity between the target and interfering sounds determined whether the significant interaction between congruency and music imagery was due to similarity between the two sounds. To do this, 18 participants rated the similarity of pairs of congruent and incongruent sounds to each other.

Each of the 48 target sounds in each category (half unfamiliar sounds) was paired with either a congruent or incongruent interference sound. 18 participants completed the similarity ratings.

Analysis of similarity ratings consisted of a three-way repeated measures ANOVA, with the factors being sound category, familiarity and congruency. This revealed a near significant main effect of sound category, *F(3, 54) = 2.984, p = 0.060, partial ŋ2 = 0.149*. Music pairings received the highest similarity ratings overall, followed by animal sounds, environmental sounds and speech. Bonferroni adjusted post hoc analysis showed a trend towards a significant difference between music and speech *(p=0.064)*. There was also a main effect of congruency, *F(1, 18) = 251.663, p <0.001, partial ŋ2 = 0.937*. Similarity ratings were higher when the target and interference sounds were congruent to each other, compared to when incongruent. There was no significant main effect of familiarity, *F(1, 18) = 1.725, p = 0.206, partial ŋ2 = 0.09*. Finally the three-way interaction between sound category, familiarity and congruency was also significant, $F(3, 54) = 3.314$, $p = 0.039$, *partial* $\eta2 = 0.163$. The analysis then focused on each familiarity level separately.

For familiar sounds there was a significant interaction between sound category and congruency, $F(3, 54) = 12.293$, $p <0.001$, *partial* $\eta2 = 0.420$. For congruent pairings, music items received higher similarity ratings, followed by speech, animal sound and environmental sound. Bonferroni adjusted post hoc comparisons revealed significant differences between music and animal sound ($p=0.013$), music and environmental sound ($p<0.001$), and a borderline difference between music and speech ($p=0.070$). There was also a significant difference between animal sound and environmental sound ($p=0.015$). There was no main effect of sound category for incongruent sounds.

The interaction between sound category and congruency was also significant, $F(3, 54) = 3.148$, $p=0.048$, *partial* $\eta2=0.156$. Again this was due to higher similarity ratings for congruent pairings. Analysis of congruent sounds only revealed a significant main effect of sound category, $F(3, 54)= 6.276$, $p=0.027$, *partial* $\eta2=0.201$, and Bonferroni adjusted post hoc analyses revealed differences between music and animal sound ($p<0.001$) and environmental sound ($p=0.022$). There were no significant effects for incongruent sounds.

Discussion

This experiment determined how perception of a sound influenced imagery vividness for familiar and unfamiliar sounds. The experiment compared imagery vividness for familiar and unfamiliar animal, environmental, music and speech sounds, like in Experiment 1, and participants rated their imagery vividness on presentation of a pictorial cue (i.e. picture of the sound item). This picture cue was used, as the previous experiment found higher ratings with picture compared to name cues. Introduction of an auditory familiarity-decision task occupied the auditory system so it could not rehearse the target sound, encouraging image generation on presentation of the visual cue. Interfering sounds could be congruent or incongruent with the target sound or white noise (used in the baseline condition).

As in Experiment 1 familiar sounds received higher imagery vividness ratings than unfamiliar sounds. The study replicated the difference between sound categories: overall speech and music items received higher imagery vividness ratings than animal and environmental sound items, but differences between sound categories were significant for unfamiliar sounds only. The fact that there was no difference between sound categories when the sounds were familiar again suggests that long-term memory for sound modulates imagery vividness, making it easier to imagine sound.

Overall target items paired with congruent interference sounds, received lower imagery vividness ratings, than those paired with incongruent interference sounds (regardless of the sound type played) or when white noise was the interference sound. There was no difference in ratings between incongruent and white noise interference conditions.

Interestingly, however, these interference effects were reliable only for music stimuli. The reason for the lack of interference effect for animal and environmental sounds may be because exemplars in these categories differ more in their acoustical properties than music. Therefore congruent interference effected ratings for music imagery may be because music items sound more similar to each other, making it harder to imagine the target item. Similarity ratings between target and interferences sounds confirmed this, revealing that for congruent sounds, similarity ratings were higher for music pairings compared to other sounds, particularly for familiar sounds. The data suggest that acoustic similarity between sounds can influence imagery vividness.

Given that music and speech consistently received higher vividness ratings than animal and environmental sounds, one suggestion is that imagery for such music and speech involves a different process or network of activation compared with the other sounds. This differing network makes imagery for these sounds more “perceptual-like” than low vividness sounds. The following experiments investigated the influence of high imaginability to investigate whether such imagery is more “perceptual-like” by looking at the effect of this vividness on sound detection.

Experiment 3: Signal Detection and Imagery Vividness

Previous research has postulated that normal vivid mental imagery lies on a continuum from more abstract to more perceptual-like experiences, with hallucinations lying at the perceptual end of this continuum (Slade & Bentall, 1988).

The precise link between imagery and perception has yet to be determined. For example, imagery can either selectively impair (Segal et al., 1970) or facilitate (Farah et al., 1983) detection of a faintly presented sound in noise, depending on the experiment conditions. In prior sound detection experiments, participants typically know the identity of the target, and the effects of imagery on detection focused on comparing performance with an explicit imagery instruction to that without. It is unclear whether participants may still create images of the target even without imagery instruction however (Stuart & Jones, 1996). Evidence for this comes from Merckelbach (2001) who stated that 32% of participants reported hearing the song ‘White Christmas’ in white noise, despite the fact no song was actually presented. This suggests that imagery of the target may be a natural occurrence in auditory detection experiments where participants know the target. Therefore the current experiment involved comparison between auditory detection under imagery conditions to a condition where the target sound was unknown, rather than relative to a ‘no imagery condition’ with a known target.

In sum, Experiment 3b examined detection of meaningful auditory sounds that varied in their familiarity and in how vivid an auditory image they evoke (Experiments 1 and 2). The study evaluated the interaction between imagery and perception, by testing sound detection when auditory imagery was prompted for a known target relative to when the target was unknown, using measures of sensitivity and response bias derived from signal detection theory.

Experiment 3b extended Experiments 1 and 2 by examining how imagery vividness and sound familiarity affect the detection of sounds in noise. Variations in familiarity between sound categories may contribute to vividness ratings however. In addition previous research has questioned the validity of vividness ratings (Allbut, 2008). Therefore Experiment 3a assessed the association between vividness and familiarity ratings, and the validity of imagery vividness ratings.

Experiment 3a. Imagery vividness and familiarity

Method

Participants

Twelve females from the University of Birmingham participated (mean age: 18.6 years old). All stated they had normal hearing in both ears and participated in exchange for credit.

Design

The experiment presented a total of 128 target sound stimuli from four sound categories. All were 22050 Hz, 16 bit sound clips, which were between 2sec in length and 3sec in length. For each sound, the application of a 20msec sound envelope controlled for the immediacy effects of each sound. The participants heard the sounds through Sennheiser HD 280 headphones, and the volume setting was the same for all participants.

Stimuli

The experiment included 128 sounds (32 per sound category, half unfamiliar). Familiar animal and environmental sounds were collected from various online sources and unfamiliar sounds were created in the same way as the previous experiments. Familiar and unfamiliar music clips also were the same as the previous experiments, however these items were played on either a piano or a clarinet. Familiar and unfamiliar speech was English and non-words spoken by either a male or a female speaker.

Procedure

The experimenter explained to the participants that the experiment would involve rating the familiarity of different sounds along with the auditory imagery vividness for these sounds (see Appendix C for instructions given to participants). In the familiarity task, participants heard a sound and immediately following this, a rating scale appeared on the screen, ranging from ‘1’ (very unfamiliar) to ‘5’ (very familiar). This indicated that participants should make their familiarity rating for the sound. In the imagery vividness task, participants heard a sound, and then a 5 sec period of white noise. The white noise was presented in order to clear sensory memory for the target item. Following this, a rating scale appeared on the screen, ranging from ‘1’ (No image at all, you only "know" that you are thinking of the object) to ‘5’ (Perfectly clear and as vivid as normal hearing).

In the first session, participants performed both the familiarity rating task and the vividness rating task. The order of the tasks was counterbalanced across participants. In the second session, participants just completed the imagery vividness task, in the same way as before.

Results

Imagery vividness ratings

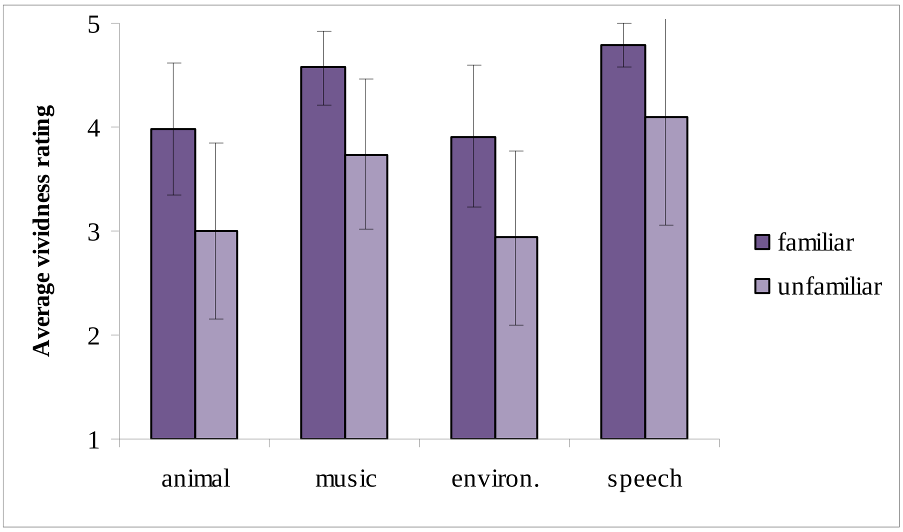


Figure 7: Average vividness rating for different familiar and unfamiliar sounds (averaged across two sessions)

Figure 7 shows the average vividness ratings for familiar and unfamiliar sounds in each sound category. Analysis of vividness ratings consisted of a three-way repeated measures ANOVA, with the factors being sound category, familiarity and session. This revealed a significant main effect of sound category, $F (3, 33) = 24.838$, $p < 0.001$, *partial* $\eta 2 = 0.693$. Bonferroni adjusted post hoc analysis revealed significant differences between animal sound and music *(p<0.001)*, animal sound and speech *(p=0.001)*, environmental sound and music *(p=0.002)*. There was also a marginally significant difference between environmental sounds and speech *(p=0.001)*. There was no significant differences between animal and environmental sounds, and between music and speech *(p>0.05).*

The significant main effect of familiarity, *F (1, 11) = 26.499, p < 0.001, partial ŋ2 = 0.707*, showed that familiar items received higher ratings than unfamiliar items. No other main effects or interactions were significant.

Vividness ratings were consistent across session, because the main effect of session was not significant. To confirm this, correlations were calculated between each condition in sessions one and two (see Table 2). All correlations were significant at the level of p<0.05, though less so for music.

Table 2: Correlation between session one and two for each sound category (r-values)

Sound category	Familiar	Unfamiliar
Animal sound	0.961**	0.805**
Music	0.578*	0.617*
Environmental sound	0.849**	0.853**
Speech	0.764*	0.896**

N= 12, * - p<0.05,-p<0.001**

Familiarity ratings

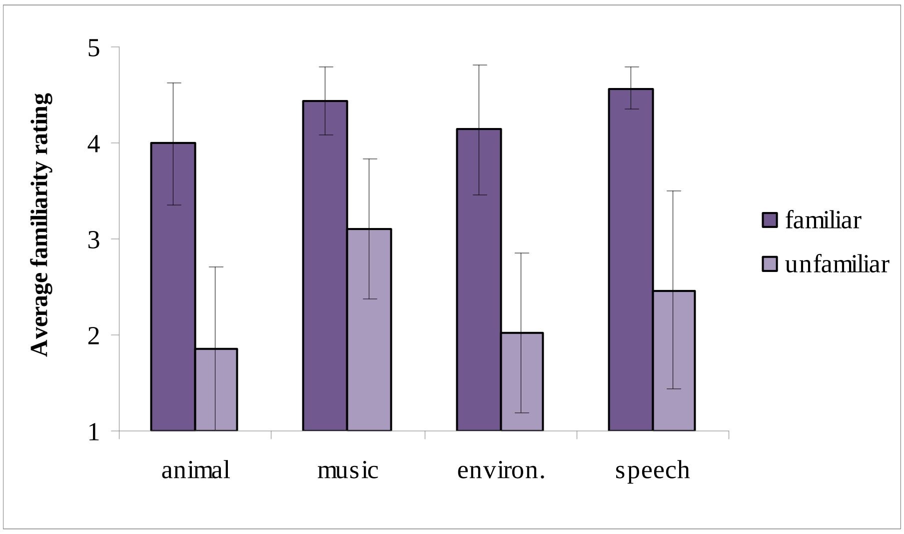


Figure 8: Average familiarity rating for different familiar and unfamiliar sounds

Figure 8 shows the average familiarity rating for familiar and unfamiliar sounds in each sound category. Analysis of familiarity ratings consisted of a 2 x 2 repeated measures ANOVA (factors: sound category and familiarity). This revealed a significant main effect of sound category, *$F(3, 33) = 7.241$, $p=0.007$, partial $\eta 2 = 0.397$*. Bonferroni adjusted post hoc analyses revealed significant differences between animal sounds and music (*$p=0.010$*). There was also a significant main effect of familiarity, *$F(1, 11) = 155.528$, $p=0.001$, partial $\eta 2=0.934$*; participants were more familiar with the familiar sounds.

There was also a significant interaction between sound category and familiarity, *$F(3, 33) = 4.153$, $p=0.023$, partial $\eta 2=0.274$*.

A separate analysis for familiar sounds revealed a significant effect of sound category, *$F(3, 33) = 5.188$, $p=0.007$, partial $\eta 2=0.320$*; speech received higher familiarity ratings than animal sounds *($p=0.025$)*.

Analysis of unfamiliar sounds also revealed a significant main effect of sound category, *$F(3, 33) = 6.595$, $p=0.010$, partial $\eta 2=0.375$*. Bonferroni adjusted post hoc analysis revealed that in this case, music received higher familiarity ratings than animal sounds *($p=0.011$)*.

Relationship between vividness and familiarity rating

A correlation revealed a significant positive association between the average vividness and familiarity ratings, *$r(128) = 0.819$, $p<0.001$*, showing that higher imagery vividness ratings were associated with higher familiarity ratings.

A two-way ANCOVA assessed the contribution of familiarity to vividness ratings, using sound category and familiarity as between participants' factors and familiarity rating as a covariate. This revealed a significant main effect of sound category, *F(3, 119) = 37.869, p<0.001, partial ŋ2=0.488*. Bonferroni adjusted post hoc analyses revealed significantly higher imagery ratings for animal compared to environmental sound *(p=0.020)*, for speech compared to animal sounds *(p<0.001)*. Also ratings were higher for music compared to environmental sound *(p<0.001)* and for speech compared to environmental sound *(p<0.001).* Finally, ratings were higher for speech compared to music *(p<0.001)* and for*)*. There was no significant main effect of familiarity, but there was a significant interaction between sound category and familiarity, *F(3, 119) = 4.544, p=0.005, partial ŋ2=0.103*. Separate analysis of familiar sounds revealed a significant main effect of sound category, *F(3, 59) = 15.295, p<0.001, partial ŋ2 = 0.437*. Again ratings were higher for music and speech than for animal sounds and environmental sounds. Bonferroni adjusted post hoc analyses revealed significant differences between animal sound and speech *(p=0.003)*, music and environmental sound *(p<0.001)* and between environmental sound and speech *(p<0.001).*

Analysis of unfamiliar sounds revealed a significant of sound category, *F(3, 59) = 36.216, p<0.001, partial ŋ2=0.648*. Bonferroni adjusted post hoc analyses revealed significant differences between animal sound and music *(p=0.024)*, between animal sound and speech *(p=0.003),* music and environmental sound *(p<0.001),* music and speech *(p=0.002)* and between environmental sound and speech *(p<0.001)*

Discussion

Experiment 3 assessed the relationship between vividness and familiarity of sounds, in the absence of other experimental manipulations. The experiment confirmed the findings of Experiments 1 and 2 - namely increased vividness ratings for music and speech items compared to animals and environmental sounds. In addition, covariate analysis confirmed that these effects were not purely due to the familiarity associated with each sound category. This analysis showed that removal of variance associated with familiarity still resulted in higher imagery vividness rating for music and speech compared to animal and speech sounds. In addition to this there was good correlation between vividness ratings across the two sessions, showing that vividness ratings are robust and consistent.

Experiment 3b examined the ability to detect the different sounds in noise. Presence of a cue (picture and sound cue vs. no cue), the imagery vividness of the stimuli (high vs. low) and the familiarity of the target sound (familiar vs. unfamiliar) were varied in this experiment. As mentioned above, Segal and Fusella (1970) found that using imagery during sound detection impaired target detection, suggesting that the image became confused with the actual target. In contrast Farah and Smith (Farah et al., 1983) found imagery use can selectively facilitate target detection, if the target and image matched. The current experiment determined which of these were true. If imagery inhibits detection, a greater bias and decreased sensitivity should be found for higher vividness sounds, particularly when cued. In contrast if imagery facilitates detection, detection should be improved by detection of higher vividness sounds and cued sounds.

Experiment 3b. Signal detection and imagery vividness

Method

Participants

40 people (35 female and 5 males, mean age 19.75 years old) from the University of Birmingham participated for psychology credit or money. All stated they had normal hearing in both ears.

Questionnaire

Participants completed the imagery questionnaire, which consisted of the new auditory imagery section and a visual imagery section from the Vividness of Visual Imagery Questionnaire (VVIQ) (Marks, 1973) (see Experiment 2 for full description).

Stimuli

This experiment presented the same 128 target sound stimuli as used in Experiment 3a, along with 5sec snippets of white noise (created using ‘Audacity’). Participants heard the sounds through Sennheiser HD 280 headphones. The sounds were presented at two different intensities in the white noise: a clearly audible level and a threshold level. The volume setting was the same for all participants

In addition, cued conditions were presented with full colour photographs. Animal and object pictures were presented with animal and environmental sounds respectively. Pictures of musical instruments were presented with music clips, and pictures of people speaking were presented with speech items.

Sounds were presented at an audible level and a threshold level during the experiment. Prior pilot testing determined the specific levels for each sound category. An average detection rate of 90 – 100% across participants defined the “audible level” and a detection rate of 70 – 75% across participants, defined the “threshold level”.

Participants detected two types of sound: those high in auditory imagery vividness (i.e. music and speech sounds) and those low in auditory imagery vividness (i.e. animal and environmental sounds; see Experiments 1 and 2). Each participant had to detect one sound category from each vividness level. Participants were randomly assigned to one of four groups: AM (Animal sounds and Music, EM (Environmental sounds and Music), AS (Animal sounds and Speech), ES (Environmental sounds and speech).

Design

This experiment had two cueing conditions. In the ‘cue’ condition, participants were cued with the target sound and associated picture, followed by a 2sec fixation period. Experiment 1 showed that cueing auditory imagery by a picture was particularly effective in evoking the image. A 5sec period of white noise followed, and the picture was presented again. Participants indicated whether the sound was present or absent in the noise, and they then had to rate how confident they were in this answer from one (not confident) to three (confident). Participants heard each sound in the white noise twice: once at an audible level, once at a threshold level. In addition on two occasions, the cue was presented, but the target was absent. There were 256 cued trials.

In the ‘no cue’ condition, the target sound and image were not presented prior to detection. Following presentation of the white noise, participants indicated whether or not a sound occurred in the white noise. They also rated their confidence in their answer, and indicated what the sound was and whether it was familiar or unfamiliar to them. Sounds were presented once at an audible level (64 trials), once at a threshold level (64 trials) and no sound was present in the remaining 128 trials. Therefore there were a total of 256 uncued trials.

In both conditions when the sound was present, it appeared in the middle portion of the 5sec white noise clip. Overall each participant completed 512 trials in a random order, across two hour-long sessions

Procedure

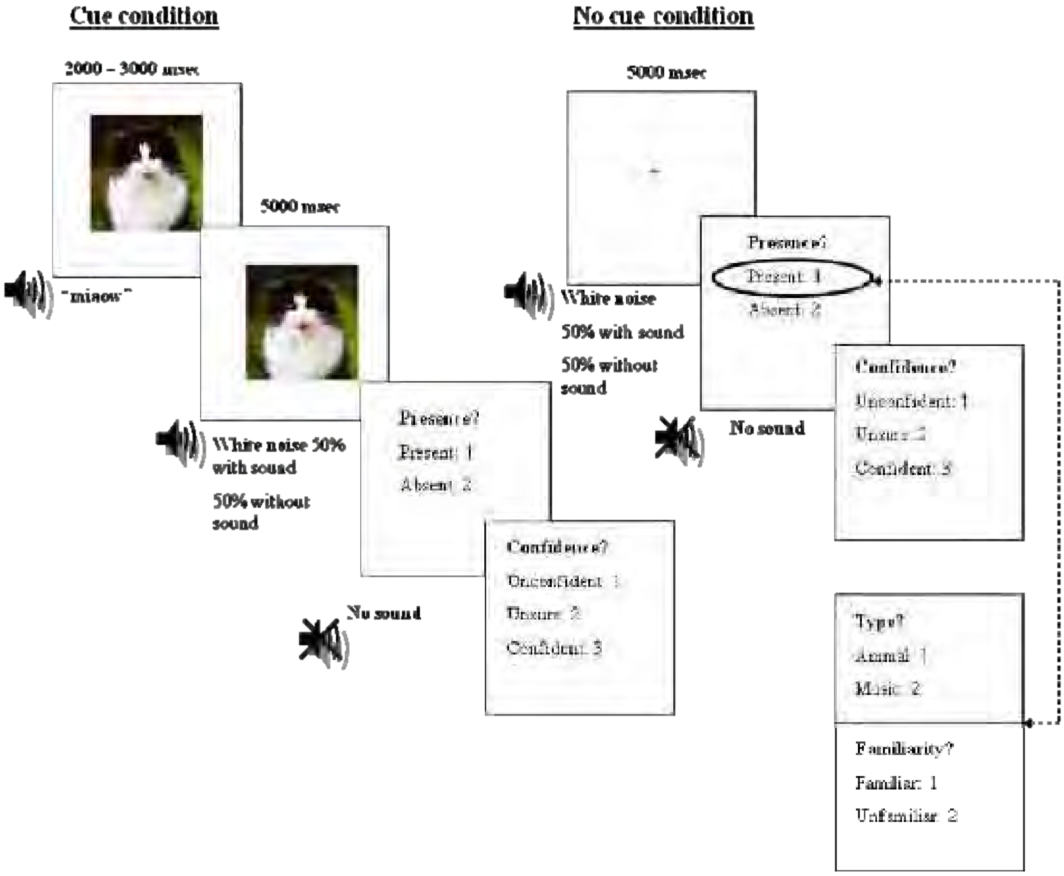


Figure 9: Experiment 3b trial procedure for each condition

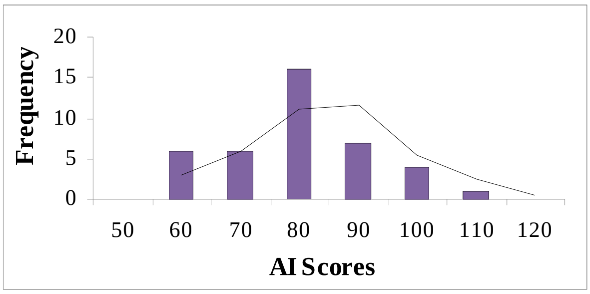
Figure 9 shows the trial procedure for this experiment in each condition. Participants were first randomly assigned to one of four groups, to determine which sounds they would detect (i.e. animal sounds and music [AM], environmental sounds and music [EM], animal sounds and speech [AS], environmental sounds and speech [ES]) (see Appendix D for instructions given to participants). The experimenter then informed the participants that the task was to detect familiar and unfamiliar sounds in white noise. They were told that the sound may or may not be presented in the white noise and that, if the sound was present, it would either be at an audible level or a threshold level. Therefore they were told they should listen very carefully to the white noise. The experimenter also informed participants that half the trials would start with presentation of a valid cue (i.e. the target sound and picture) or would receive no cue.

Notification of the type of condition appeared on the screen at the beginning of each trial (i.e. “cue” or “no cue”).

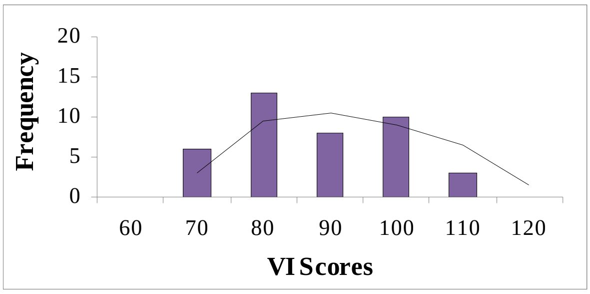
In the experiment, participants first indicated whether the target sound was present or absent in the white noise and then how confident they were in their answer, by rating from one (confident) to three (unconfident). In the ‘no cue’ condition, if they said the sound was present, participants also indicated what the sound was and whether it was familiar or unfamiliar to them. Prior to the experimental trials there was a practice session of four trials, after which any questions the participant’s had were answered and then the experimental trials began. Each of the two sessions participants took part in consisted of eight blocks, with an opportunity for a break after each block. Half the participants completed an imagery questionnaire at the beginning of the first session, and half before the second session.

Results

Imagery Questionnaire



a)



b)

Figure 10: a) Distribution of Auditory Imagery scores; b) Distribution of Visual Imagery scores (VVIQ)

The average auditory imagery rating was 75.00 (SD =12.01) and the average visual imagery rating was 82.78 (SD = 12.21) (see Figure 10 for distribution of these scores).

Sound detection task

Two analyses assessed performance in the sound detection task[2]. The first analysis focused on the sound detection measures of response bias (c) and sensitivity (d') and the confidence in responses. This determined how sound detection is affected by characteristics of the sounds (vividness and familiarity) and availability of a cue.

[2]Appendix H contains further analyses of the acoustic characteristic of target sounds and how these characteristics influence sound detection.

The second analysis focused on correlations with the signal detection measures. Firstly it determined participant's reality testing abilities, by looking at the associations between accuracy and confidence rating. Secondly it determined the direct association between vividness, familiarity rating and signal detection measures.

For each condition, data from the four vividness group pairings were combined and the proportion of hits and false positives for each imagery vividness and familiarity level were calculated. The hit rate was calculated from the total proportion of trials where the participants correctly stated that the target was present, at the threshold level. The false alarm rate was calculated from the total proportion of trials where the participant incorrectly stated that the target was present when it was absent.

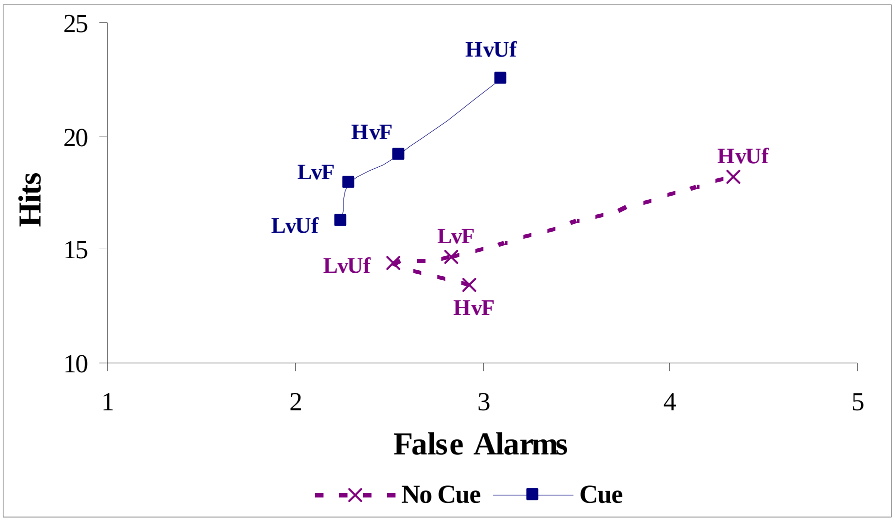


Figure 11: Hits and FA for each condition

Figure 11 shows a plot of hits against false alarms for each condition in the sound detection task. This shows that detection was generally better in the cue condition, as indicated by the higher hit rate and lower false alarm rate. In addition high vividness sounds had the highest hit and false alarm rates in each condition (particularly the unfamiliar sounds).

Analysis One: The effect of vividness, familiarity and cue condition on performance on sound detection task

Signal Detection Analyses

Because the main measure of this task required participants to make a two alternative forced choice (i.e. target presence vs. target absence) analysis of the data utilised signal detection theory.

This theory posits that when a participant needs to choose between two alternatives, two factors contribute to the decision. The criterion reflects the amount of response bias that a person has towards responding positively in ambiguous environments (i.e. sound plus noise conditions). A variety of factors can influence this response bias, such as familiarity or relevance of the stimuli (Stanislaw & Todorov, 1999). These factors lead the participant to adopt either a conservative or liberal criterion, so that greater or lesser amounts of signal are required for the participant to answer positively. β and *c* are the most commonly used measures of criterion, though *c* is most appropriate for inferential statistical analyses (Pastore, Crawley, Berens, & Skelly, 2003). A higher value of the *c* indicates a more conservative response bias.

Sensitivity (d') is the other factor that can influence the detection of a target. This is a measure of perceptual sensitivity to the target which can also be affected by a number of factors including intensity or brightness. Higher values of this measure indicate greater sensitivity.

In this study, the measures of *c* and d' were calculated using the hit and false alarm rates reported above. These signal detection measures were calculated using the excel equations reported in Stanislaw et al. (1999). Here *c* was used as the criterion measure calculated, because this measure is most appropriate for analysis using inferential statistics.

Criterion

Table 3: Average criterion in each condition (standard deviation in brackets)

	Low Vividness		**High Vividness**	
	Familiar	**Unfamiliar**	**Familiar**	**Unfamiliar**
Cue	1.63 (0.39)	1.69 (0.43)	1.59 (0.42)	1.48(0.38)
No Cue	1.75 (0.44)	1.81 (0.53)	1.79 (0.54)	1.52 (0.47)

Table 3 shows that for familiar sounds, high and low vividness sounds did not differ. For unfamiliar sounds however, the criterion was lower for high vividness sounds than low vividness sounds.

Analysis consisted of a three-way repeated measures ANOVA (factors: vividness, familiarity and condition). This revealed a significant main effect of vividness, *F(1, 39) = 5.919, p = 0.019, partial n2 = 0.133,* showing a greater bias for high vividness sounds. The main effect of condition was also significant, *F(1 ,39) =6.916, p = 0.012, partial n2 = 0.151,* revealing a greater bias in the cue condition. The main effect of familiarity was not significant, however.

There was a significant interaction between vividness and familiarity, $F(1, 39) = 15.104, p < 0.001, \text{partial } \eta^2 = 0.279$. High and low vividness sounds did not differ when the sounds were familiar, *t(39) =0.088, p=0.930*. When the sounds were unfamiliar, high vividness sounds had a significantly lower criterion than low vividness sounds, *t(39) = 3.807, p<0.001*. This did not vary as a function of the cue conditions. No other interactions were significant.

d'

d' is a measure of sensitivity to a signal. A high d' prime indicates a higher sensitivity to the target sound.

Table 4: Average d' in each condition (standard deviation in brackets).

	Low Vividness		High Vividness	
	Familiar	**Unfamiliar**	**Familiar**	**Unfamiliar**
Cue	1.42 (0.75)	1.42 (0.77)	1.44 (0.72)	1.45 (0.74)
No Cue	1.33 (0.80)	1.34 (0.88)	1.40 (0.89)	1.18 (0.76)

Table 4 shows the average d' for each condition. Analysis of the d' data consisted of a three-way repeated measures ANOVA (factors: vividness, familiarity and condition). This revealed no significant main effects or interactions.

Confidence Ratings

Table 5: Average confidence rating in each condition (standard deviations in brackets)

	Low Vividness		High Vividness	
	Familiar	**Unfamiliar**	**Familiar**	**Unfamiliar**
Cue	2.77 (0.15)	2.73 (0.22)	2.81 (0.17)	2.83 (0.17)
No Cue	2.61 (0.23)	2.52 (0.25)	2.45 (0.37)	2.66 (0.35)

Table 6 shows the average confidence rating in hits in each condition. Analysis of confidence in hits consisted of a three-way repeated measures ANOVA (factors: vividness, familiarity and condition). The main effect of condition was significant, *F*

(1, 39) = 80.089, p < 0.001, partial η^2 *= 0.673*. Participants were less confident that they heard a sound in the no cue condition. There was a significant interaction between vividness, familiarity and condition, *F (1, 39) = 8.714, p = 0.005, partial* η^2 *= 0.183*.

For familiar items, the interaction between vividness and condition was significant, *F(1,39) = 10.433, p = 0.003 , partial* η*2 = 0.211*. There was no significant difference between low and high vividness sounds in the cue condition, *t(39) = 1.415, 0.165*, but in the no cue condition, participants were significantly more confident that they heard the low vividness sounds, *t(39) = 2.831, p=0.007.*

For unfamiliar items, the interaction between vividness and condition was not significant.

The confidence ratings in false positives could not be analysed because in many conditions participants did not make any false positives responses.

Analysis Two: Correlational analysis

The correlational analysis assessed reality monitoring differences between the cue and no cue conditions, and the association between sound detection performance and vividness and familiarity. Because of the number of comparison made here, the stringent criteria of p<0.01 was set for the analysis to be considered significant.

Correlations between hits, false positives and confidence ratings

Hits, false positives and confidence ratings were calculated over items (i.e. for each specific sound) and correlated as a measure of participant's accuracy in judging their internal state. False alarm rate and confidence ratings could not be calculated for the no cue condition because false alarms were not made to specific sounds in this condition, rather to a general sound category (i.e. high or low vividness sound).

Table 6 shows correlations between hits and the confidence ratings in each condition, as well as the false positives from the cue condition and the associated confidence ratings.

Table 6: Correlations between confidence ratings and hit and false positive rates.[3]

Measure		Correlation
Hits	**Cue**	0.863**
	No Cue	0.736**
False Positives	**Cue**	-0.046

**** = p<0.001**

The significant correlations between hits and confidence ratings showed an association between a higher hit rate and increased confidence in a sounds presence, which was greatest when participants received a cue.

Correlation with vividness and familiarity

This analysis assessed the association between these vividness, familiarity and signal detection measures, to determine the direct influence of vividness and familiarity on performance.

[3] Some sounds received either no hits or no false positives therefore these correlations may differ in the number of items involved in each comparison.

The sound detection measures (c, d' and confidence ratings) were calculated over item for each target sound, and correlated with average vividness and familiarity rating for each sound.

None of the correlations reached significance at the criteria of $p<0.01$. A near significant correlation showed an association between (i) confidence on hit responses in the cue condition and (ii) vividness rating, $r\ (126) = 0.206$, $p = 0.021$, however. Participants were more confident in hearing sounds that were rated as more vivid.

During the current study, participants also completed an imagery questionnaire. Correlations were computed between scores on this questionnaire and overall average scores of each sound detection measure. This determined whether individual differences in imagery abilities affected sound detection.

Correlations between the average auditory imagery rating and the average c, d′ and confidence rating, were not significant ($p>0.01$).

Summary

The ANOVA investigations of sound detection tasks revealed significant effects of vividness, familiarity and condition on bias towards responding and on confidence in an item’s presence, but no affect on sensitivity to target items. Participants showed a greater bias to reporting the presence of high vividness, unfamiliar sounds and low vividness, familiar sounds in noise. Confidence ratings mirrored this affect, showing that participants had greater confidence in the presence of such sounds, particularly in the cue condition. In addition both the bias and confidence ratings were higher in the cue condition compared to the no cue condition.

The correlational analysis confirmed that the presence of the cue aided judgment of internal state during sound detection, evidenced by the greater association between high rate and confidence rating in the cue compared to no cue condition.

In addition correlation with vividness and familiarity ratings showed an association between greater confidence in hit responses and higher imagery vividness. No relation was found between performance on the sound detection task and vividness or familiarity ratings however.

Discussion

Experiment 3b investigated whether cueing, imagery vividness and familiarity can affect sound detection. Analysis of the d′ data revealed no difference in sensitivity dependent on the vividness of the sounds, the familiarity of the sounds, or cue presence. In contrast, analysis of the criterion, revealed a stronger bias towards high vividness sounds and to sounds in the cue condition. The significant interaction between vividness and familiarity on the response bias, revealed that for familiar sounds, there was no difference in bias between high vividness (i.e. music and speech) and low vividness sounds (i.e. animal and environmental sounds), but for unfamiliar sounds, there was a greater bias for high vividness stimuli (i.e. music and speech).

These data suggest that the detection of auditory signals in noise depends on biases based on the kinds of learned knowledge participants have concerning particular sounds. With familiar sounds, participants adopt a moderate response criterion irrespective of how vivid the image of the sound is. The response bias does not increase because the stimulus is already associated with a vivid auditory image; familiarity protects against this bias. In contrast, for unfamiliar items the response bias increases for stimuli that are associated with higher auditory imagery vividness.

This finding could be explained in terms of variations in the cognitive effort required during the detection of different types of sounds. For instance because music and speech items are high vividness items, they require little cognitive effort for the generation of auditory images. This is especially so when the item is familiar, such that the image is too defined for it to be confused for an actual percept, making false positive responses less likely to occur. When items are unfamiliar however, though little cognitive effort is required to image the high vividness item, the lack of familiarity means the image is ill-defined, leading participants to adopt a more liberal criterion (and so, more false positive responses).

With low vividness sounds however, greater cognitive effort is required to create the image. This is particularly effortful for unfamiliar items, such that more effort is allocated to image generation, rather than to sound detection, making false positive responses less likely. For familiar items, less effort may be required (as the item is more defined), making false positive responses more likely to occur. In this explanation therefore, detection bias is determined by the balance between cognitive effort and vividness of the sound: if the sound is very vivid less effort is required to imagine the item therefore the bias decreases. Similarly if the sound is associated with weak vividness it becomes too effortful to produce an image, leading to adoption of a more conservative bias.

The study aimed to provide support for one of two theories about the detection of sounds following imagery. Segal and Fusella (1970) suggested that imagery and detection compete for resources, resulting in lower sensitivity to sounds when imagery was used during detection. In contrast Farah and Smith (1983) found improved sensitivity for sounds when imagery was used, suggesting that imagery facilitates detection through attentional recruitment, so long as the target and image matched.

Analysis of the criterion confirmed that the cue resulted in participants adopting a more liberal criterion however, as the response bias was greater in this condition compared to the no cue condition. This did not serve to improve sensitivity. Therefore the study suggests that provision of a cue did not lead participants to confuse their image for the target item, or prime detection of the target item - contrary to both Segal and Fusella (1970) and Farah and Smith (1983). Provision of the cue did increase confidence in the presence of the target item however, and the study also revealed a stronger correlation between hits and confidence ratings in the cue condition compared to the no cue condition. Therefore the cue appeared to exert a mild facilitatory effect on target processing. Provision of an imagery period upon presentation of the cue, prior to detection of the target item, may have resulted in a stronger effect of cue.

Vividness, familiarity and condition also influenced confidence ratings. Analysis revealed higher confidence ratings for cued items. In addition there was a significant interaction between all three factors. For familiar items, there was no difference between low and high vividness sounds in the cue condition, but, in the no cue condition, participants were significantly more confident that they heard the high vividness sounds. Participants' knowledge of what a familiar, vivid sound should be may increase their confidence, because of a greater specification of the 'template' to which they match these sounds, compared to low vividness sounds. Under noisy presentation conditions there may always be some mismatch between the stimulus and the template, but for high vividness sounds, participants may tolerate this mismatch more. In a sense, the better specified the template for the sound, the less cautious the participant. In contrast, for unfamiliar items, participants were more confident in hearing high vividness sounds than low vividness sounds across both cue conditions. There may be a poorer specification of the template for unfamiliar items, however. In this case having a more vivid associated auditory image may generate a decreased response criterion and more confident responding.

General Discussion

The current study investigated auditory imagery vividness for different categories of familiar and unfamiliar sounds, and how vividness variations influence sound detection.

Experiment 1 assessed the effect of different cues on the introspective vividness of auditory images. The self-rated vividness of auditory imagery was higher for pictures compared to name cues across sound categories, suggesting that pictures evoke auditory images more clearly than words. Previous research has not investigated the affects of different cues on the vividness of auditory imagery, though there has been relevant work conducted on visual and auditory memory (Lehmann & Murray, 2005).

Experiment 2 investigated how a secondary task affects auditory imagery for different sound categories. Following presentation of the target, participants judged the familiarity of a second sound, and then rated their imagery vividness for the target. The study hypothesised that familiarity judgements of similar sounds would disrupt auditory imagery for sounds– perhaps akin to interference in verbal short-term memory when items have overlapping phonological representations (e.g. Baddeley et al., 2000). Consistent with this argument, congruent interfering sounds (matching the imaged stimuli category) decreased the vividness ratings for auditory images. Further investigation revealed that congruent interference decreased imagery vividness of music items only, suggesting that acoustic similarity between interference and target sounds can affect imagery vividness, rather than their semantic association. Assessment of similarity between target, congruent and incongruent sounds confirmed this: congruent pairings of music clips received higher similarity ratings than other congruent pairings. Therefore the lack of congruency effect on other sound categories may be because exemplars in these sound categories vary more in their acoustic characteristics, so that there is less interference from using common representations in imagery and memory processing.

In both Experiments 1 and 2, music and speech sounds received higher ratings than animal and environmental sounds, and familiar sounds received higher ratings than unfamiliar sounds. A possible explanation for the variations in imagery vividness between sound categories is that in our everyday lives we more commonly generate images of music and speech than animal and environmental sounds, making these types of images more robust. For instance Bailes (2006) used an experience sampling method and found that people report imagining music 35% of the times that they were sampled. Therefore usage of music imagery appears to occur frequently throughout the day, and this frequency of usage may correlate with imagery vividness, though further research would determine this.

Another possibility for the difference in vividness between sound categories is that music and speech are just more familiar than animal and environmental sounds. Experiment 3a found that the differences in imagery vividness between sound categories were independent of the rated familiarity of the stimuli, since introducing familiarity as a covariate in the analysis of vividness ratings did not eliminate the differences between the sound categories. The ability of participants to subvocalise some but not other sounds (music and speech but not environmental and animal sounds) may explain the variation in imagery vividness between sound categories. Reisberg et al. (1989) described the distinction between 'pure' auditory imagery and 'enacted' auditory imagery. Pure auditory imagery uses the phonological loop and the central executive to generate an exact image of the sound, whereas enacted imagery also requires the use of subvocalisation to imagine sounds. It follows that environmental and animal sounds are pure images as it is difficult to subvocalise such sounds (Dick et al., 2007), whereas music and speech are enacted images because rehearsal of these sounds involves subvocalisation. In addition Baddeley and Andrade (2000) found a reduction in imagery vividness for music and speech following articulatory suppression. This last result may also explain why the difference between familiar and unfamiliar sounds was smaller for music and speech items than between animal and environmental sounds, since these participants could use subvocalisation to rehearse and enhance imagery for such sounds. Further research could investigate the effects of articulatory suppression on the ability to generate auditory imagery for different sounds. Specifically further studies could evaluate whether articulatory suppression has a selective effect on music and speech imagery compared to animal and environmental sound imagery, which would support Reisbergs' theory of auditory imagery generation for different sounds.

Experiment 3a also assessed the association between vividness and familiarity. This revealed good correlation between vividness and familiarity ratings. In addition the study confirmed the reliability of the auditory vividness ratings, by finding good correlation between participants vividness ratings made on two separate occasions. Therefore this study shows that imagery ratings are robust and independent of familiarity, adding strength to the theory that they are accurately measuring auditory imagery abilities.

In the vividness rating experiments, familiar sounds received higher ratings across all sound categories compared to unfamiliar sounds, suggesting that prior experience with sounds is important for imagery clarity. This supports and extends Baddeley and Andrade's (2000) finding of the differences in imagery vividness between meaningful and nonsense stimuli. In their experiment meaningful stimuli consisted of spoken sentences, whilst nonsense stimuli consisted of mixed up portions of spoken sentences. Participants' ratings of imagery vividness were lower for nonsense than for meaningful items. Here we extend this result to show that there are differences in auditory imagery across a range of known items that still vary in their familiarity.

Experiment 3b determined whether imagery inhibits or facilitates sound detection. Sensitivity (d′) to the target did not differ between the cue and no cue conditions, suggesting that prior knowledge of the target items identity neither facilitated nor interfered with its detection. The cue did influence the bias (criterion) however, so that participants more biased to reporting the sounds presence, following a cue. In addition, the bias was greater for high vividness sounds, showing that vividness of a sound can modify the response criterion people adopt, encouraging people to have a more liberal bias for such items. This factor did not influence sensitivity however, suggesting that subjective vividness of imagery for a sound does not affect sensitivity to the target. Analysis of the criterion data also revealed an interaction between vividness and familiarity, showing a greater bias to familiar compared to unfamiliar sounds when the sound was low vividness, but a greater bias to unfamiliar sounds that were high in vividness. This may reflect different decision criteria between low and high vividness items which varies as a function of their familiarity. With low familiarity items there is a bias to accept a sound that is high in vividness. High familiarity, however, seems protective against this bias, so that participants do not differ in their bias between low and high vivid items. One account of this is that participants have a more finely tuned template for highly familiar items, making it less likely that they will accept a match to that template.

Analysis of confidence ratings given on hit trials revealed an interaction between vividness, familiarity and condition. For familiar items, low and high vividness sounds did not differ significantly in the cue condition, but in the no cue condition, participants were significantly more confident that they heard the high vividness sounds. Analysis of unfamiliar sounds revealed greater confidence when sounds were low vividness and for sounds in the cue condition, but the vividness and condition did not interact. This interaction firstly suggests that in the cue condition, highly familiar sounds override the effect of vividness on confidence: if a sound is familiar to the participants they are more confident in its presence, regardless of the sound category. In the no cue condition however, information about the targets identity is unavailable, therefore in this case, participants appear to use a ‘template’ for high vivid items, and are more confident in the presence of these items when they appear.

Segal et al. (1970) also investigated the effect of sound familiarity on sound detection and reported poorer detection when participants imaged an unfamiliar sound compared to a familiar sound. They suggested that the image of the unfamiliar sound interfered with the detection of the tone. This result runs parallel to the effects of familiarity on the detection of stimuli whose auditory images are low in vividness. This may be because imagining the sound required extra processing capacity to aid detection, leaving little left for actual detection.

Finally the validity of vividness of imagery measures requires comment. The current study used a vividness measure to investigate imagery abilities in response to different types of sounds and different task requirements. Such a measure requires participants to rate the strength of their auditory imagery for different sounds. It is therefore likely that individual differences in criteria for the ratings arise, according to what individual participants categorise as a strong auditory image. Nevertheless the data show that vividness of imagery is a robust measure, is replicable in individual participants, is independent of variations in sound familiarity across categories, and it effects detection of sounds, particularly effecting decision confidence. In addition, previous research has revealed an association between self-rated visual imagery vividness and visual cortex activation (Cui et al., 2007). Such research into auditory imagery is lacking, therefore a future study should investigate the association between self-rated auditory imagery vividness for different sounds and levels of neural activation. This would offer evidence that vividness ratings are a true representation of actual imagery abilities.

Conclusion

The current study found differences in subjective ratings of imagery vividness for different types of sounds, dependent on sound familiarity, cues given to imagine a target sound and interference from similar sounds. In addition the study found that vividness ratings are reliable and are not a product of the familiarity associated with the sounds. Experiment 3 hypothesised that highly vivid auditory images and provision of a valid cue would interfere with detection of faintly presented sounds in noise, due to confusion between the image of the item and the actual percept.

High vividness sounds, and cued sounds increased participant's response bias and confidence, and vividness interacted with familiarity to affect these measures also. Sensitivity to the target sounds was unaffected however.

In addition the study revealed that vividness, familiarity and condition affected confidence ratings. These last results may link to the differential specificity of auditory templates for the sounds associated with highly familiar and less familiar stimuli, and to how the vividness of a sound interacts with the stored template. This template varies according to the familiarity associated with the sounds, so that when unfamiliar, there is an increased bias towards accepting that a sound high vividness. Increased familiarity however protects against this bias, so that there is no difference between low and high vividness items. Therefore high familiarity may lead to development of a more defined template which participants are less likely to accept a match to that template.

Chapter Three. The neural correlates of animal and environmental sound perception and imagery

Abstract

The development of more advanced neuroimaging techniques has made imaging of auditory imagery and perception possible (Specht et al., 2003; Bunzeck et al., 2005; Hall et al., 1999; see Chapter 1). The current study employed an event related sparse sampling technique to investigate the relations between auditory imagery and perception for animal and environmental sounds. The participants underwent an fMRI brain scan, whilst they either imagined or listened to sounds in silence and rated the vividness or prototypicality of the item respectively. These conditions were compared to a baseline condition where participants either did nothing or made a button press in response to a written number cue. Contrasting the perception condition with baseline resulted in bilateral activation of the superior temporal gyri (STG), whereas there was no increased neural response for the imagery condition when compared with the baseline. More important, right STG showed larger responses to animal sounds than environmental sound both during the perception and imagery trials. Further analysis focused on the imagery condition. Regions within the frontal lobe were associated with auditory imagery vividness ratings across participants. Psychophysical interaction (PPI) analysis revealed increased in coupling between the frontal regions and the right STG during the imagery condition. Taken together the data indicate that the right STG distinguishes between acoustically similar but semantically different non-verbal sounds (animal sounds vs. environmental sounds), while frontal areas (and frontal-STG connectivity) link to differences in the vividness of auditory imagery.

Experiment 4. A sparse sampling analysis of animal and environmental sound perception and imagery

Introduction

Prior to recent advances in brain imaging, the study of the brain mechanisms of auditory imagery has been relatively limited (Jensen, 2005). Recent analyses however have revealed that auditory imagery activates similar regions to that of perception (Yoo, Lee, & Choi, 2001; Bunzeck et al., 2005), that activation of auditory areas occurs during auditory hallucinations (Dierks et al., 1999; Bentaleb, Beauregard, Liddle, & Stip, 2002) and that damage to auditory regions can disrupt auditory imagery (Zatorre & Halpern, 1993).

Neuropsychological studies classically posit left hemisphere regions as the location of auditory speech perception, mainly due to the propensity for speech comprehension disorders following left temporal lobe damage (Boatman, 2006). A number of fMRI studies of speech perception (Specht et al., 2003; Binder, Frost, Hammeke, Rao, & Cox, 1996) and speech imagery (Shergill et al., 2001; McGuire et al., 1996a; Scott et al., 2000) supported this. This is a contentious issue however, as other studies find bilateral temporal lobe activation during speech perception tasks, rather than just left hemisphere activity (Binder, 2000).

In contrast to speech sounds, previous studies suggest that imagery and perception of non-verbal sounds (i.e. music, animal and environmental sounds) are more lateralised to the right temporal lobe. For instance Zatorre and Halpern (1993) showed that patients with right hemisphere damage had impaired performance on imagery and perception pitch comparison tasks, compared to the left hemisphere damaged participants. Specht et al. (2003) found that perception of tones, animal sounds and music resulted in stronger activation in the right superior temporal sulcus.

The majority of auditory imagery and perception studies have focused on music and speech. Few studies have investigated other non-verbal, non-musical sounds, such as animal and environmental sounds. Therefore comparison of the neural correlates of such sounds will clarify the organisation of the auditory system and determine whether the neural organisation for different auditory categories has a semantic basis or whether it is purely acoustic. This is the aim of the current study. Previous studies have investigated the neural correlates of animal and environmental sounds, but have focused on either perception or imagery.

Studies of animal and environmental sound perception

Lewis et al. (2005) found bilateral activation of the middle superior temporal gyrus in response to animal sounds, and stronger, more distributed activity in the left hemisphere in response to perception of tool sounds. Other studies have also found similar regions of activation in response to animal and tool sounds (Altmann, Doehrmann, & Kaiser, 2007; Doehrmann, Naumer, Volz, Kaiser, & Altmann, 2008). It is unclear whether this dissociation is due to the semantic differences between the categories, or acoustic properties however, because the sound categories differed in their spectrograms and harmonic to noise ratios (Lewis et al., 2005).

As well as differences in which brain regions are activated by animal and environmental sounds, studies also reveal temporal differences in neural responding. Murray, Camen, Andino, Bovet and Clarke (2006) studied the temporal time course of brain activation to animal and tool sounds, using auditory evoked potentials. Unlike the Lewis et al. (2005) study, Murray et al. (2006) matched their sound stimuli for their acoustic characteristics. The study revealed stronger activation for tool sounds in the right temporal lobe and left inferior frontal gyrus, 70 ms following stimuli onset. Over the 155-257 ms time period, activity in the bilateral temporal and premotor cortices peaked later and longer for animal sounds. Therefore this study shows early differences in activity between living and non-living sounds can occur, which are independent of the acoustic differences between these sound categories. Studies investigating the neural localisation of differences between living and non-living sounds using fMRI is lacking however.

Studies of animal and environmental sound imagery

Studies into the neural correlates of animal sound and environmental sound imagery are rare. Bunzeck et al. (2005) found an association between environmental sound perception and bilateral superior temporal lobe activity (including the primary and associative auditory cortices). In addition activation associated with environmental sound imagery overlapped the activation in the bilateral associative auditory cortices, but not the primary auditory cortices. This suggests that, as for neuroimaging studies of music (Zatorre et al., 1993; Halpern et al., 2004) and speech (Shergill et al., 2001; McGuire et al., 1996a) environmental sound imagery overlaps some of the regions involved in sound perception.

Kraut (2005) investigated memory for non-verbal sounds by comparing neural activation in response to threatening and non-threatening animal and environmental sounds using fMRI. Participants indicated whether the sounds were real or non-real (non-real sounds were scrambled versions of real sounds). Contrasts between threatening and non-threatening sounds revealed activation in the right superior temporal and inferior frontal gyri, and the right parietal lobe. More relevant to the current study however, comparison between animal and environmental sounds revealed activation in the STG bilaterally and in regions in the left frontal lobe, while contrasts in the other direction revealed no significant differences. The stronger activation in auditory areas for animals compared to environmental sounds, and for threatening compared to non-threatening stimuli, may reflect the evolutionary significance of animal sounds (supported by activation of the same STG area during threatening sounds). Studies of auditory imagery and perception for animal and environmental sounds therefore revealed dissociations in activations between categories, even when controlling for acoustic differences, but an overlap in regions involved in imagery and perception.

Though an important method for determining neural activation to auditory stimuli, fMRI investigation is extremely noisy (between 95 – 140 dB) because of vibrations caused by interactions between the gradient coils and the magnet (Hall, Goncalves, Summerfield, Foster, & Bowtell, 2000). This makes auditory processing experiments difficult. Though the amplitude can be reduced with earplugs and ear defenders or headphones, it cannot be reduced to a level where it can no longer be heard (Ravicz et al., 1999). Therefore it may be difficult for participants to hear the experimental stimuli, which may distract attention away from the task (Gaab, Gabrieli, & Glover, 2007a). Previous studies argued that as scanner noise is present in all conditions, the effects on the activation are cancelled out by contrasting the experimental conditions with a control condition (Gaab et al., 2007a; Belin, Zatorre, Hoge, Evans, & Pike, 1999). Gaab et al's. (2007b) study however, revealed different intensities of activation when scanner noise was added to auditory stimuli compared to that during scanner noise alone or auditory stimuli alone, suggesting that the scanner noise masks some of the activation in response to the auditory stimuli. ‘Sparse temporal sampling’ techniques reduce the impact scanner noise has on performance by introducing silent gaps in the scanning sequence, to allow for stimulus presentation, or task completion.

Hall et al. (1999) compared continuous imaging to a sparse imaging technique, using continuous speech stimuli. The study found activation in the same regions for both methods, but analysis revealed greater a signal to noise ratio when using the sparse sampling technique. Similarly Gaab et al. (2007a) compared two sparse sampling techniques to an event related design. In the auditory task participants heard four words, and indicated which two were the same. The sparse sampling designs collected one volume either every 16sec (jittering the trial presentation within this period), or every 6sec at the end of each trial. The event related design acquired volumes every 2sec, resulting in continuous scanner noise. All three designs revealed significant bilateral temporal lobe activity. Both sparse temporal sampling designs revealed greater signal intensity in Heschl’s gyrus compared to the event related design however. This therefore suggests that the scanner noise may mask some of the intensity of the stimulus-induced signal in continuous scanning procedures.

A key aim of sparse sampling designs is to reduce the influence of scanner noise on activation to auditory stimuli. Therefore particular attention should be paid to the time course and peak levels of activation of the hemodynamic response function (HRF) of the auditory cortices, to ensure that the scanner noise does not confound with activation. Hickok et al. (1997) found that the response to auditorily presented words and non words started 3sec post-stimulus and peaked at 5-6sec post stimulus, which is similar to that found in previous studies of visual regions (Bandettini, 1995, cited by Hickok, 1997) and auditory studies (Hall et al., 2000).Therefore when designing a sparse sampling experiment, it may be beneficial to sample activation at different time points, to ensure that the peak of activation is encompassed.

The current study employed a novel clustered acquisition technique fMRI to investigate auditory imagery and perception of sounds. This reduced the possibility of the scanner noise interfering with ability to perform the tasks, and also avoided the scanner noise masking the activation to the target stimuli. The experiment itself investigated the neural responses to animal and environmental sounds, the neural overlap between auditory imagery and perception, and how subjective ratings of auditory imagery vividness affect neural activation. In the visual domain, Cui, Jeter, Yang, Montague and Eagleman (2007) found that participants who rated themselves as having higher visual imagery also had stronger activation in the visual cortex during a visual imagery task. Similar studies are lacking the auditory modality. The current study hypothesised that similar effects would occur in the auditory domain, showing an association between higher vividness rating and increased neural response. Finally, previous research suggested increase coupling between sensory and frontal cortices during visual imagery compared to perception (Mechelli, Price, Friston, & Ishai, 2004).Therefore the study also aimed to test for change in coupling between auditory and frontal regions during imagery compared with the perception condition. Investigation of the relationship between auditory imagery and perception will provide much needed expansion of the functional organisation of the auditory system, and will determine the association between auditory imagery vividness and neural activation.

Method

Participants

13 right-handed healthy volunteers participated in this experiment (2 males and 11 females, 24.5 years). All had normal hearing and no previous history of neurological or psychiatric disorders.

Stimuli

The stimuli were 12 animal sounds and 12 environmental sounds. All sounds were 2sec in length. The environmental sounds were vehicles, household objects and tools.

The animal and environmental sounds were matched for their root mean square intensity (RMS) and harmonic-to-noise ratio (HNR). One way ANOVA's revealed that these measures did not differ significantly (RMS: $F(1, 22) = 0.025, p=0.967$, *partial* $\eta2<0.001$; HNR: $F(1,22)=1.385, p=0.252$, *partial* $\eta2=0.059$). The average pitch (in Hz.) also did not differ between the two categories, $F(1,22)=0.232, p=0.635$, *partial* $\eta2=0.010$.

A pilot study obtained familiarity ratings for each of the sounds (ratings from one to five, one being least familiar), from ten people who did not participate in the fMRI experiment. This revealed that the two sound categories did not differ in rated familiarity, $F(1, 22) = 0.622, p = 0.439$, *partial* $\eta 2 = 0.028$.

Design

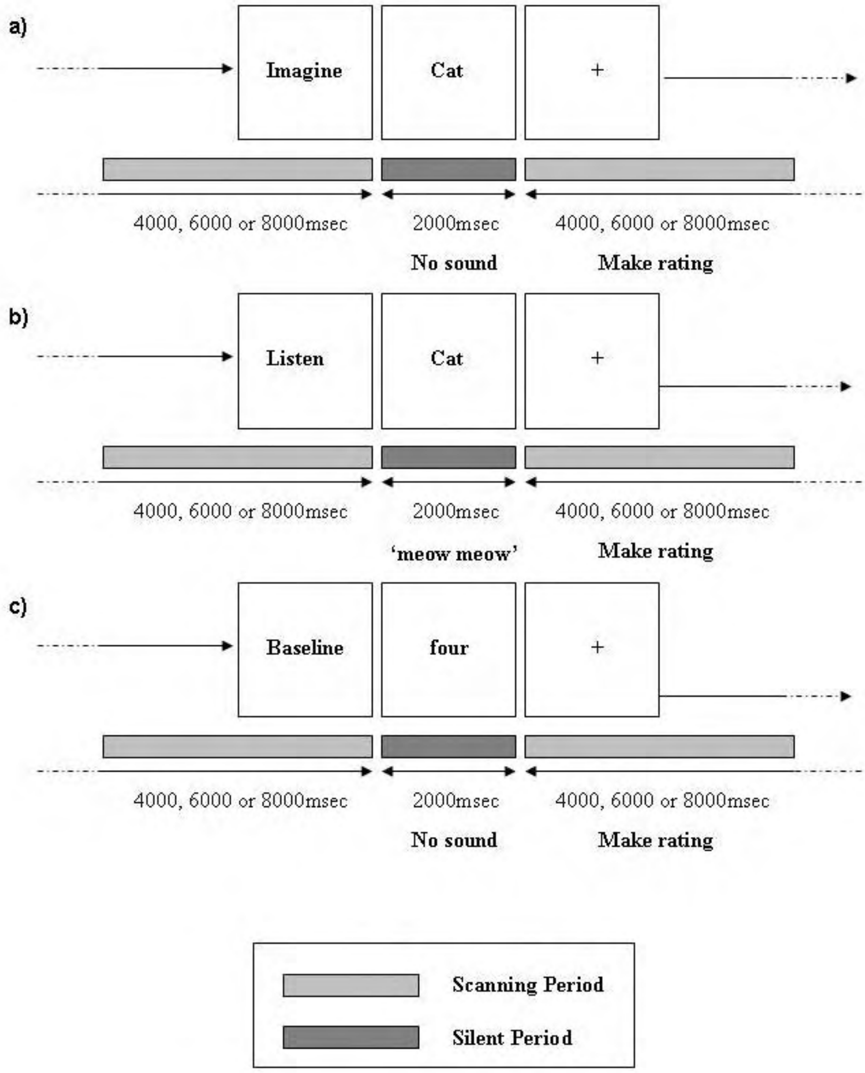


Figure 12: Procedure of three main conditions: a) imagery condition b) perception condition c) baseline condition

The study examined the relationship between condition (imagery vs. perception) and category (animal vs. environmental sound), in comparison to baseline (see Figure 12).

In the imagery conditions participants first received a cue for the upcoming condition on the screen (i.e. “imagine”) for 1sec, followed by the name of an animal or environmental sound for 2sec. During this presentation time, participants imagined the sound, and upon presentation of a fixation cross they rated the intensity of their auditory imagery for that item, on a four-point scale (‘1’corresponding to “no image” and ‘4’ corresponding to “image is perfectly clear”).

In the perception condition, participants also received a 1sec condition cue (i.e. “listen”). They then saw the name of an animal or environmental sound and heard the sound associated with this item, for 2sec. During this presentation time, participants listened to the sound, and then upon presentation of the fixation cross, rated how prototypical the sound was of the named item, on a four-point scale (‘1’corresponding to “not at all prototypical”, and ‘4’ corresponding to “very prototypical”).

In the baseline condition participants received a condition cue (i.e. “baseline”) and then a written number (one, two, three or four) for 2sec. When the fixation cross appeared, they pressed the button associated with the written number. This was in order to equate for the demands of word reading and motor response in the experimental conditions. In addition, 15% of the trials were null events in which participants did not receive a stimulus. These trials added additional jitter within the event related design, to facilitate the haemodynamic responses estimation for the different events.

Importantly, in all the trials during the 2sec period that followed the 1sec cue of each event (imagery, perception and baseline) scanner data acquisition was paused, to minimise interference from scanner noise (see below for details). 81 trials were presented overall for the imagery condition and for the perception condition (36 animal and 36 environmental sounds, 9 random animal or environmental sound trials). For the baseline condition, 36 trials were presented in which participants made a button response. The experiment was divided into 12 sessions. In each session the conditions (five conditions plus null events) were presented randomly in blocks of 18 stimuli. Each sound was presented three times in the imagery condition and three times in the perception condition, with the distance between repeating trials maximised.

Stimulus presentation and response collection were performed using E-prime (Psychology Software Tools, Pittsburgh, PA). Auditory stimuli were presented using MR compatible electrostatic headphones. Due to technical problem, two participants could only hear the perceptual trials through the right headphone, and intermittently through the left headphone, however the analysis still included their data since auditory stimuli presented in both sides of space are represented in both hemispheres. Participants also wore earplugs, and the volume was set at a comfortable level by playing the experimental stimuli prior to starting the experiment. Before participating in the fMRI experiment, participants listened to the sounds to familiarise them with the stimuli and to ensure that they could imagine them easily. Participants also completed a block of the experiment, to familiarise them with the procedure, as it would be in the scanner (white noise simulated the scanner noise). This also familiarised the participants with the task of rating their own auditory imagery.

fMRI data acquisition

Functional MR images were collected using a 3-T Achieva Phillips scanner and an eight channel phase array coil. Functional scans were blood oxygenated level dependent (BOLD) contrast weighted acquired using echo planar imaging (EPI). Thirty-four axial slices were acquired in each volume, with a slice thickness of 2mm and a gap of 1mm, positioned so that most of the brain was covered, including the whole of the temporal and frontal lobes. Repetition time (TR) was 2sec, TE was 35msec, 80° flip angle and in plane resolution 3mm x 3mm.

To achieve an optimal environment to facilitate imagery and perception of auditory stimuli, a sparse sampling technique was employed. In this procedure silent periods were introduced at the times when sounds had to be imagined (imagery condition) or listened to (perception condition). This method also ensured that effects of imagining or listening to the stimuli would not be confounded by activation due to scanner noise. Sparse sampling was achieved by manually programming the scanner to pause for a whole TR (2s) every 4, 6 or 8sec. The acquisition time was randomly distributed and this manually programmed sequence was repeated for each block (see Figure 12 for design). Stimuli were presented during these silent periods in a pseudo-random order. This sparse sampling method insured that the fMRI data were acquired when the BOLD response peaked, 3-6s after the event occurred (Friston et al., 2003; Hall et al., 1999; Gaab et al., 2007a; Amaro et al., 2002).

fMRI Data Analysis

The fMRI data were analysed using SPM5 (Wellcome Department of Imaging Neuroscience, London; www.fil.ion.ucl.ac.uk/spm).

Pre-processing: The functional images were realigned to the first image (Ashburner & Friston, 2003) and unwrapped to account for movement by distortion interactions (Andersson, Hutton, Ashburner, Turner, & Friston, 2001).The data were normalised by transformation to standard MNI space (Ashburner et al., 2003). Finally the data were smoothed using an $8mm^3$ FWHM Gaussian kernel to account for residual inter-subject variability.

Statistical analyses: Statistical analysis was performed in two steps. First the effect size for each condition was estimated for each subject using the general linear model framework (GLM). To achieve this, a model was created in which the onsets of each of the five conditions of interest (i.e. two perception conditions, two imagery conditions and the baseline condition) were defined. These onsets were convolved with the canonical HRF (Friston, Harrison, & Penny, 2003). In addition for each of the experimental conditions, we added the ratings of imagery vividness and prototypicality as covariates. An additional two nuisance covariates were added to account for T1 1^{st} and 2^{nd} order saturation effects following each pause of the scanner. In order to combine the model and the fMRI data, the time window where fMRI data were not collected (i.e. during the 2sec silent pauses) were removed from the model. To correct for signal changes due to head movement, the 6 realignment parameters were included in the design matrix. Finally, an additional set of harmonic regressors was used to account for any temporal low-pass frequency variance within the data with a cut-off of 1/128Hz.

Voxel Based Analysis

In order to assess the activation to each condition while controlling for effect of visual stimulation and response production, the experimental conditions were contrasted with the baseline condition for each participant. The resultant contrasted images were entered into a new factorial design model in which participants were treated as random factor. This model did not assume independent or equable variance. Finally, a further analysis was carried out on regions within the auditory cortex. Results are reported using threshold of $p<0.001$ uncorrected (unless specified otherwise). WFU_pickatlas was used to identify the primary and associative auditory Brodmann areas. Tests for differential effects within these regions were performed using SPSS16.

Ratings analysis

An analysis also determined where activation was modulated by the imagery vividness ratings, to assess the regions that were sensitive to the ratings in the imagery conditions.

Psychophysical interaction (PPI) analysis

Psychophysical interaction (PPI) analysis (Friston et al., 1997) investigated the change in connectivity between different brain regions as a function of task. This analysis measures the change in coupling between two regions in response to a psychological variable. In the current study this analysis focused on testing whether imagery is associated with a different network architecture than perception. Previous research has shown increase coupling between associative sensory and frontal cortices during visual imagery compared with visual perception conditions (Mechelli et al., 2004).

Here the analysis investigated whether similar effects will be observed in the auditory cortex. Change coupling with the auditory associative cortex during auditory imagery vs. perception were measured using PPI with the seed region defined as the right auditory cortex (MNI coordinates: 68 -24 10) based on the group results (see below). For each participant, the analysis used an eigenvector depicting the time series from a 10mm sphere around the seed region in the right auditory association cortex. The PPI was computed for the imagery vs. perception contrast. A new model was estimated for each participant, containing the regressors from the previous analysis, a regressor of the region of interest time course and the PPI regressor (psychological x physiological interaction: imagery vs. perception x by the seed region response).

Results

Behavioural performance

The differences between animal sounds and environmental sounds for imagery ratings and prototypicality ratings were analysed separately. One participant did not make any ratings in either the imagery or perception condition and was not included in analyses.

Imagery ratings were analysed using a paired samples t-test. This revealed no significant differences between the two sound categories, $t(11)=2.054$, $p=0.065$ though there was a trend for higher imagery ratings for animal sounds. For the prototypicality ratings animal sounds were considered significantly more representative of the sounds they were supposed to be, $t(11)= 2.623, p=0.024$.

The average imagery vividness and prototypicality ratings for each participant were also correlated for each sound category. This revealed no significant correlation across participants between the two ratings for animal sounds $r(12) = 0.131, p=0.686$ or environmental sounds $r(12) = 0.336, p = 0.285$.

Voxel Based Analysis

Table 7: Peak activations for each contrast (threshold of p<0.001 uncorrected, unless stated otherwise)

					Peak coordinates		
Contrast	**L/R**	**Region**	**BA**	**T**	**x**	**Y**	**z**
		Perception vs. Baseline					
	Left	STG	41	7.21**	-50	-32	10
	Left	STG	42	6.79**	-60	-32	10
	Left	STG	22	6.34**	-48	-18	-4
	Right	STG	42	6.72**	68	-22	8
	Right	STG	22	6.62**	68	-24	0
	Right	STG	22	6.10**	56	-18	2
	Left	IFG	46	4.23**	-48	30	18
	Left	IFG	46	4.00	-48	32	10
	Left	IFG	46	3.59	-50	42	10
	Left	SMA	6	3.95	-2	20	48
Environmental Sound Perception vs. Animal Sound Perception							
				No significant voxels			
Animal Sound Perception vs. Environmental Sound Perception							
	Right	STG	22	3.42	62	-2	-6
	Left	STG	22	3.17	-60	-8	-2
		Imagery vs. Baseline					
				No significant voxels			
Environmental Sound Imagery vs. Animal Sound Imagery							
				No significant voxels			
Animal Sound Imagery vs. Environmental Sound Imagery							
AI vs. EI	Right	STG	42	2.75*	68	-24	10
	Right	STG	22	2.74*	60	-2	-4

**** = p < 0.05, FWE-corrected; *p<0.005 (uncorrected). STG = superior temporal gyrus, IFG = inferior frontal gyrus, SMA = supplementary motor area**

Perception Condition vs. Baseline

Contrast between the perception condition and the baseline condition, revealed the largest regions of activation in the bilateral temporal lobes (incorporating both the primary and associative auditory cortices), the left inferior frontal gyrus and the left supplementary motor area (see Figure 13). Table 7 shows the coordinates of the peak areas of activation.

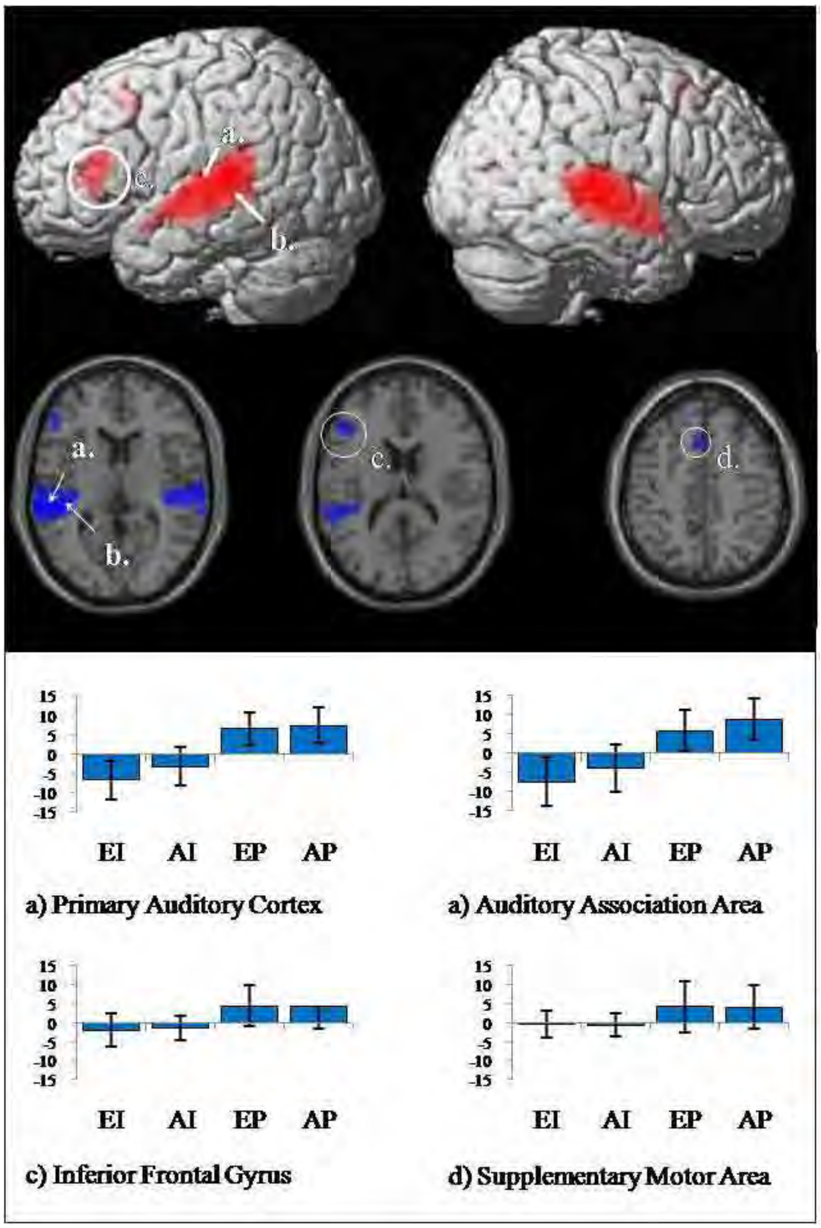


Figure 13: Perception vs. Baseline
Top row: SPMs ($p<0.001$ uncorrected) overlaid on a lateral view of a single subject rendered brain; Middle row: the same SPMs overlaid on a single subject axial T1 contrast images. In all images, the left hemisphere is on the left side. Bottom row: Beta values for the peak areas of activation in a) Primary Auditory Cortex, b) Auditory Association Area[4] c) Inferior Frontal Gyrus and d) Supplementary Motor area (regions circled on the SPM images above)

[4] Beta plots for the primary auditory cortex and auditory association area are of the left hemisphere.

Imagery condition vs. baseline

Whole brain analysis of this contrast revealed no significant voxels (*p>0.001* uncorrected) for the imagery condition.

Animal sounds vs. environmental sounds

The environmental sound vs. animal sound contrast revealed no above threshold activations. Contrasting animal sounds with environmental sounds revealed significant bilateral activation in the superior temporal lobes (see Figure 14). Table 7 shows the peak coordinates for these regions.

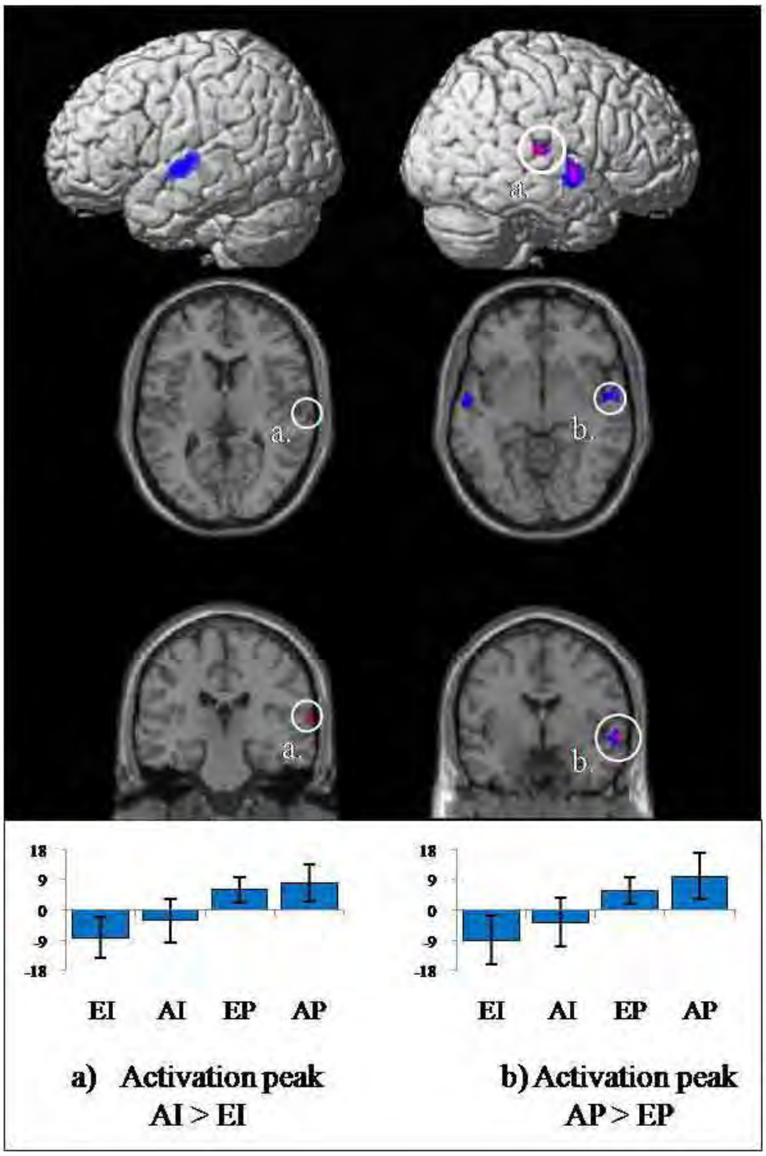


Figure 14: Animal sounds vs. environmental sounds

Top row: SPMs of animal > environmental overlaid on a lateral view of a single subject rendered brain (for illustration purposes the threshold here is p<0.005 uncorrected, cluster >50 voxels); Middle rows: the same SPMs overlaid on a single subject axial and coronal T1 contrast images. In all images, left hemisphere is presented on the left side; Red = imagery, blue = perception, purple = overlap; Bottom row: Beta plots for the peak activations for a) AI>EI and b) AP>EP (peak regions of activation circled on the SPM images, see table x for coordinates of the peaks)

Animal sound imagery– baseline vs. environmental sound imagery - baseline

Similarly to the perception condition, contrasts between environmental sound imagery and animal sound imagery revealed no above threshold responses (*p>0.001* uncorrected). Contrasting between animal sounds and environmental sounds, there was significant right superior temporal lobe activation for the imagery task (see Figure 14). Table 7 show the peak coordinates for these regions.

Region of Interest Analyses

The perception condition revealed considerable activation in the bilateral superior temporal lobes, incorporating both the primary and associative auditory regions. In order to investigate neural activation in these regions in more detail, a region of interest (ROI) analyses was conducted. The primary auditory cortex incorporates the Heschl's gyrus, and corresponds with BA 41, whilst the associative auditory cortex incorporates the planum temporale and corresponds with BA 42 and BA 22. These regions were masked using the PickAtlas tool and the peak coordinates in each region were determined from the perception vs. baseline contrast. A 10mm volume of interest mask was then specified surrounding the peak activation for each region and hemisphere, and the beta values were extracted. Figure 13 shows the average beta values for the primary auditory cortex, auditory association area, left middle frontal gyrus and the left supplementary motor area

Activation in the bilateral Primary Auditory Cortex (PAC), and Auditory Association Area regions were analysed a four way split-plot ANOVAs. The factors were sound type (animal vs. environmental sounds), task (imagery vs. perception), hemisphere side (left vs. right) and region (primary vs. associative auditory cortices). This analysis revealed a significant main effect of sound type $F(1, 48) = 9.229, p = 0.004,$ *partial* $\eta 2 = 0.161$; activation was significantly stronger for animal sounds than environmental sounds. There was also a significant main effect of task $F(1, 48) = 273.116, p <0.001,$ *partial* $\eta 2 = 850$; activation was significantly stronger for the perception task compared to the imagery task. The main effects of side and region were both not significant *(side:* $F(1, 48) = 0.0395, p = 0.843,$ *partial* $\eta 2 < 0.001;$

region: $F(1, 48) = 0.036$, $p = 0.850$, partial $\eta 2 < 0.001$). There were no significant interactions.

Correlations with behavioural ratings

Prototypicality ratings for each sound category were correlated with beta values of the primary and associative auditory cortices during the perception conditions. These analyses revealed no significant correlations between prototypicality ratings and activation strength in either region or hemisphere ($p>0.001$).

Imagery vividness ratings for each sound category were also correlated using a similar procedure, which revealed no significant correlations between vividness ratings and activation strength in either region or hemisphere ($p>0.001$).

Left Inferior Frontal Gyrus (IFG)

The left IFG showed a significant cluster of activation for the contrast between the perception and baseline condition (see Table 7), and the beta values from a 10mm sphere around the peak were analysed. Figure 13 shows the average beta values for each sound category in each condition. A significant main effect of task was found, $F(1, 12) = 11.500$, $p = 0.005$, *partial* $\eta 2 = 0.489$: activation was stronger during the perception conditions compared to the imagery conditions. The main effect of sound category was not significant, $F(1, 12) = 0.047$, $p = 0.832$, *partial* $\eta 2 = 0.004$, nor was the interaction between sound category and task, $F(1, 12) = 0.130$, $p = 0.725$, *partial* $\eta 2 = 0.011$.

Left Supplementary Motor Area (SMA)

The left SMA also showed significant activation in the contrast between the perception and baseline conditions and a region of interest analysis was conducted on the beta values from a 10mm sphere around the peak. Figure 13 shows the average beta values for each sound category, in each condition. There was a trend towards greater activation during the perception task compared to the imagery task, $F(1, 12) = 4.646$, $p = 0.052$, partial $\eta 2 = 0.279$. No significant effect of sound category or interaction between sound category and task was found ($F(1, 12) = 0.157$, $p = 0.699$, partial $\eta 2 = 0.013$ and $F(1, 12) = 0.002$, $p = 0.968$, partial $\eta 2 < 0.001$, respectively).

Negative Beta Values

All the ROI analyses resulted in negative beta values for the imagery conditions relative to baseline. To investigate these in greater detail, the opposite contrast was calculated (i.e. baseline vs. imagery conditions; see Figure 15). This showed a large amount of significant activation in the bilateral superior temporal lobes (see Table 8 for peak regions of activation).

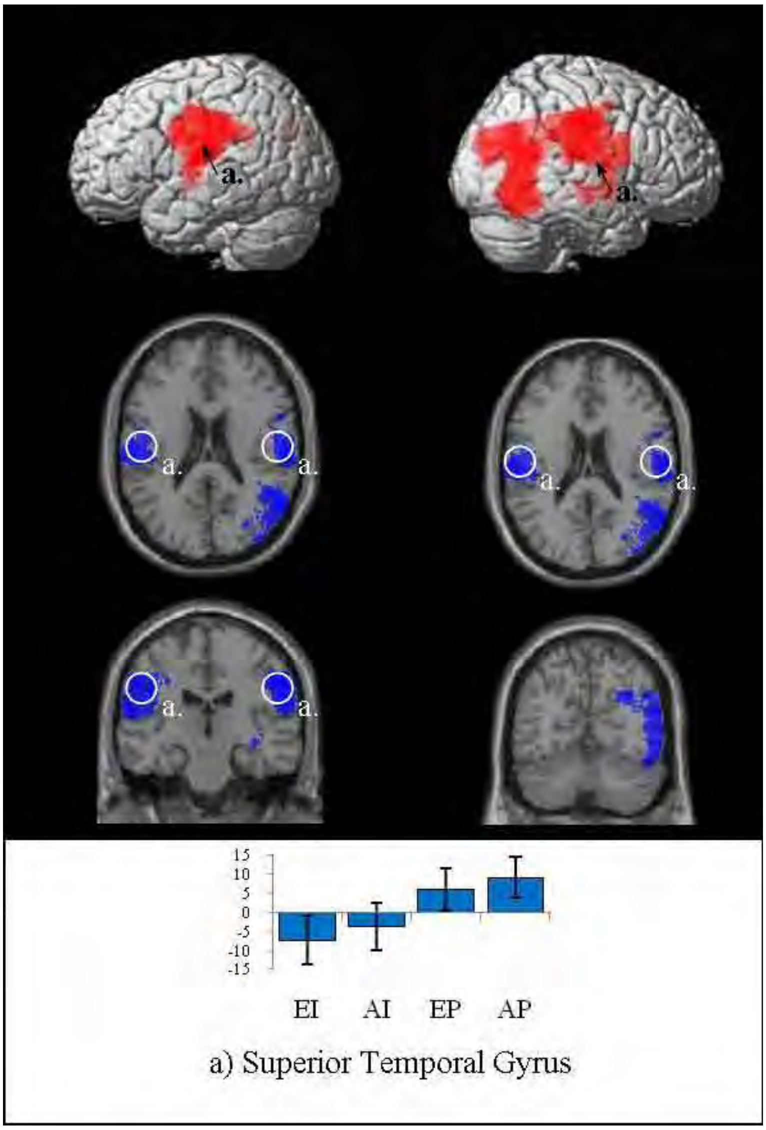


Figure 15: Deactivation in the imagery conditions (Baseline vs. Imagery)

Top row: SPMs of animals > environmental overlaid on a lateral view of a single subject rendered brain (for illustration purposes the threshold here is $p<0.0001$ uncorrected); Middle rows: the same SPMs overlaid on a single subject axial and coronal T1 contrast images. In all images, left hemisphere is presented on the left side; Bottom Row: Beta plots of peak areas of activation.

Table 8: Peak regions of deactivation during imagery conditions (baseline versus imagery) (p<0.001 uncorrected)

				Peak coordinates		
L/R	**Region**	**BA**	**t**	**x**	**Y**	**z**
Right	STG	22	5.48	56	-16	38
Right	STS	22	5.39	60	-60	24
Left	STG	22	4.92	-46	-24	32
Left	STG	22	4.67	-58	-18	18
Right	Lateral Occipital Gyrus	18	4.88	36	-94	-4
Right	Cingulate Gyrus	24	4.59	2	-30	44
Right	IFG	45	4.39	26	28	26
Left	Cingulate Gyrus	18	4.3	-12	-2	50
Left	MTG	21	4.26	-60	-60	-8
Left	MFG	9	4.08	-4	50	-8
Left	Precuneus		4.02	-24	-66	34
Right	Cuneus		3.97	12	-82	2
Right	MFG	9	3.91	8	50	28
Right	MTG	21	3.84	62	-38	-12
Left	Cingulate Gyrus	18	3.82	-8	-2	-2
Left	MFG	9	3.74	-12	50	12
Left	STG	22	3.73	-26	-50	24
Left	Insula	13	3.71	-42	-2	-2
Left	Cingulate Gyrus	18	3.62	-12	-52	12
Right	MFG	9	3.60	48	26	24

Rating Analyses

In order to determine the areas that were sensitive to vividness ratings in the imagery condition, a separate analysis was conducted. This determined the extent of the modulation of neural response by subjective imagery vividness. This analysis revealed a number of significant clusters (see Table 9, Figure 16) most notably in the left middle frontal gyrus (MFG) and the bilateral insula lobes.

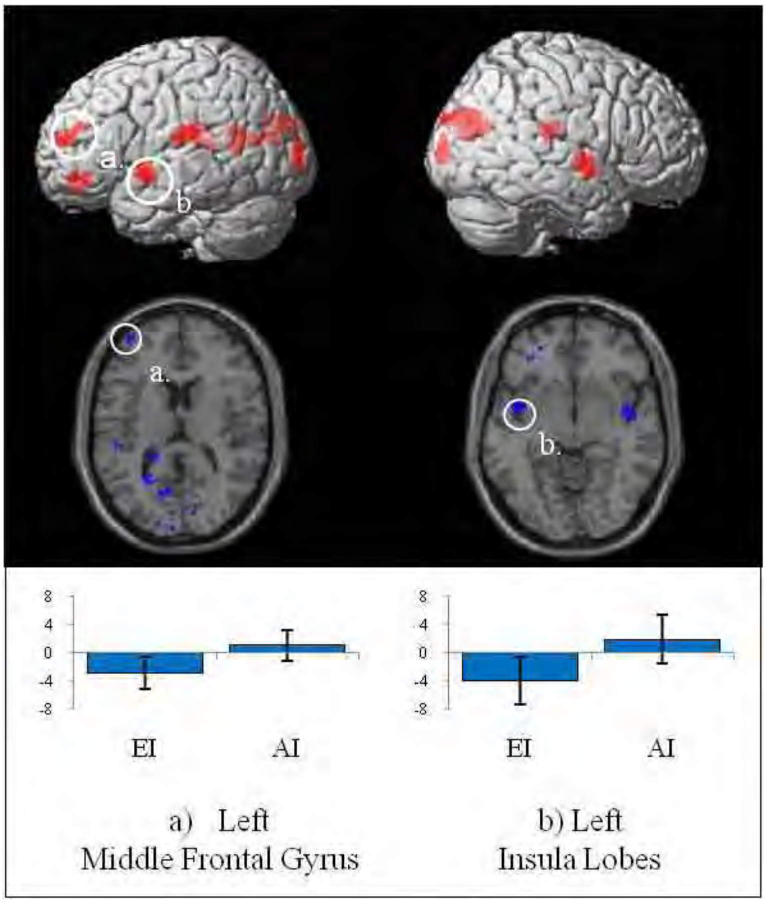


Figure 16: Rating analysis - environmental sound imagery and animal sound imagery
Top row: SPMs of effects of vividness ratings on imagery responses (p<0.001 uncorrected) overlaid on a lateral view of a single subject rendered brain; Middle row: the same SPMs overlaid on a single subject axial T1 contrast images. In all images, left hemisphere is presented on the left side; Bottom row: Beta plots for peak activations in the rating analysis a) Left MFG, b) Left insula lobe (see Table 9 for coordinates of peaks).

Table 9: Peak activations for imagery ratings (p<0.001 uncorrected)

					Peak Coordinates		
Hemisphere	**Anatomical Structure**	**BA**	**Cluster Size**	**F**	**X**	**y**	**z**
Left	MFG	10	8	12.1	-40	54	14
Left	Insula lobe	52	9	10.16	-42	2	-6
Right	Insula lobe	52	5	10.15	46	-4	-4

Psychophysical Interaction (PPI) Analysis

This analysis investigated changes in coupling with the auditory cortex during imagery compared to perception trials. The seed region was in the right auditory cortex (MNI coordinates: 68, -24, 10). Voxel-based analyses showed activation in this region in the perception vs. baseline contrast and in the animal sound imagery vs. environmental sound imagery contrast (see above).

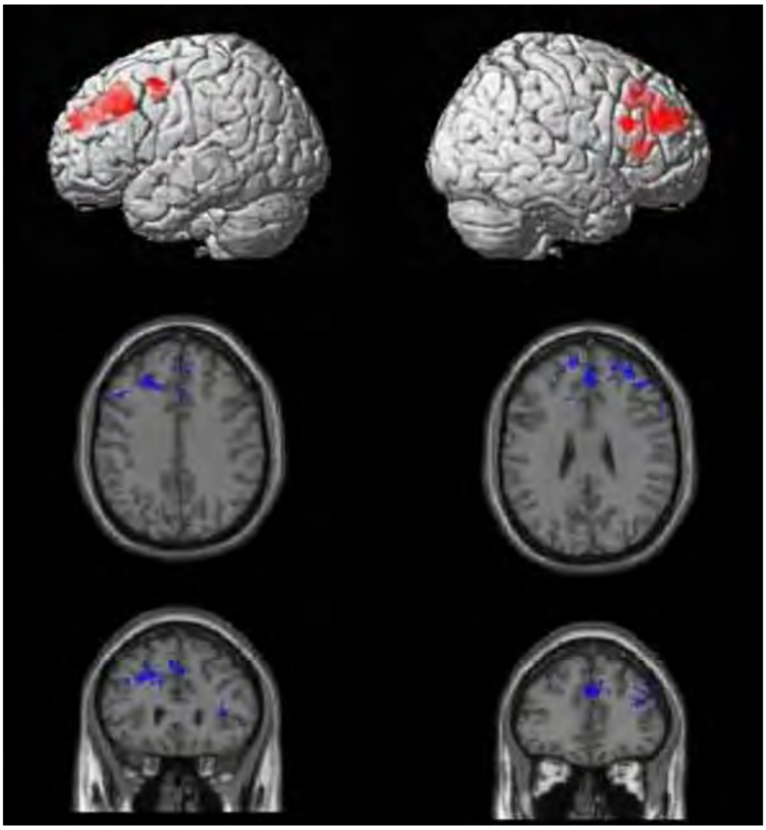


Figure 17: PPI analysis - increased coupling with the right auditory cortex
Top row: SPMs showing increase coupling with right auditory cortex during imagery trials overlaid on a lateral view of a single subject rendered brain(for illustration purposes the threshold here is p<0.005 uncorrected, cluster >50 voxels); Middle and Bottom rows: the same SPMs overlaid on a single subject axial and coronal T1 contrast images. In all images, left hemisphere is presented on the left side.

Table 10: Peak activation in PPI analysis of imagery vs. perception contrast (p<0.001 uncorrected)

				Peak coordinates		
L/R	**Region**	**BA**	**T**	**x**	**y**	**z**
Right	MFG	9	6.31	42	36	28
Right	IFG	45	5.98	54	20	24
Left	SFG	10	5.72	-10	52	26
Left	MFG	9	5.18	-24	30	38
Left	MFG	9	5.11	-38	4	52
Right	MFG	10	5.03	36	30	10
Right	MFG	10	4.72	30	44	24
Left	SMA	6	4.56	-4	32	42
Left	SFG	9	4.47	2	42	30
Right	ACC	32	4.43	6	22	40

This PPI analysis revealed increase coupling of the right secondary auditory cortex and bilateral frontal lobes, particularly with the middle and superior frontal gyri during imagery vs. perception conditions (see Figure 17 for activation pattern and Table 10 for peak voxels). The opposite contrast (perception vs. imagery) revealed no significant results.

Discussion

Experiment 4 showed (i) dissociations between the semantically (but not acoustically) different categories of animal sounds and environmental sounds, (ii) auditory imagery and perception activate similar regions of the auditory cortex, (iii) subjective ratings of auditory imagery intensity are associated with activation in the frontal lobes, and iv) there is significant connectivity between auditory cortices and frontal regions, during imagery. The implications of these findings are discussed below.

Auditory perception

Relative to the baseline condition, the perception of environmental and animal sounds activated the primary and secondary auditory cortices bilaterally; in addition there was reliable activation in the left SMA and IFG. Previous studies of auditory imagery also found SMA activation, and may reflect subvocalisation during such imagery (McGuire et al., 1996a). Activation in this region has not been documented during auditory perception before, however. It is unlikely that the SMA activation reflects subvocalisation of word reading here, since the control condition also required reading, as well as the experimental conditions. An alternative explanation is that listening to animal or environmental sound evokes other cross-modal associations such as motor imagery associated with the use of objects (Lewis et al., 2005) or for animals in motion, which would involve activation of motor-related areas such as the SMA (Lambert, Sampaio, Scheiber, & Mauss, 2002).

Previous studies also implicate the inferior frontal gyrus (IFG) in auditory imagery, particularly in association with inner speech processes (McGuire et al., 1996a; Shergill et al., 2001). Other studies have found left IFG activation during speech perception (Specht et al., 2003) and tool sound perception (Lewis et al., 2005), however. Unfortunately, the paucity of prior research on imagery and perception for such non-verbal sounds hampers interpretation of the activity in these regions. Further research into activation associated with such stimuli, in comparison to speech is required to determine the nature of activation in these regions.

Region of interest analysis found stronger activation strength for animal sounds compared to environmental sounds. The contrast between animal sounds and environmental sounds confirmed this, revealing significant bilateral activation in the primary auditory cortices (stronger in the right hemisphere). Contrasts in the opposite direction (environmental sounds versus animal sounds) revealed no significant activations. This pattern of findings is similar to that of Kraut et al. (2006). Their study employed a continuous fMRI procedure to investigate memory for animal sounds, in comparison to that of tool sounds. This revealed significant right STG activation when contrasting animal sounds with tool sounds, but no significant activation for the opposite contrast. Kraut et al. (2006) attributed their findings to animal and environmental sounds forming distinct semantic groups. Kraut et al. (2006) did not match the sound categories for their acoustic properties (most notably the harmonic to noise ratio and frequency) therefore the contrast in acoustic properties could have contributed to the difference. In the current study the acoustic properties between the two sound categories were matched for sound intensity, harmonic to noise ratio and sound frequency, yet the differences in neural activation still emerged. This suggests that there is a non-acoustic basis for the neural dissociation between the categories. As suggested by Kraut et al. (2006), the contrast may reflect a semantic difference between animal and non-animate environmental sounds. This would fit with both imaging data and neuropsychological studies showing that double dissociations can exist between knowledge of animals and objects (Forde, Francis, Riddoch, Rumiati, & Humphreys, 1997; Derenzi & Lucchelli, 1994). The sensory/functional account of semantic deficits suggests that dissociations may be because animals are more strongly defined by visual-perceptual than functional properties (e.g., they often have fur, have tails, walk on four legs etc), whereas the opposite holds for inanimate objects (Warrington et al., 1984) . The greater activation for animal over environmental sounds here may reflect the stronger weighting of perceptual features (in this case auditory features) in the identification of animate sounds. There are other possibilities however. For instance, the auditory representations of different animals may be relatively similar (e.g., barks, yelps, snorts) compared with the similarity between environmental sounds, and this perceptual similarity may lead to greater neural activity in regions associated with auditory processing (see Forde & Humphreys, 2001, for arguments about perceptual similarity in the visual modality). Indeed in Gygi et al.'s (2007) multidimensional scaling investigation of the similarity between environmental sounds, vocalisations formed the largest similarity cluster. The authors attributed this to the fact that such sounds are generally more harmonic and more similar to each other in acoustic properties compared to other categories of environmental sounds.

Additionally, the dissociation may also reflect that, here, participants rated the environmental sounds as less prototypical than animal sounds. This reduced prototypicality may have resulted in reduced levels of activation. Against this last point however, correlations between prototypicality ratings and activation strength in primary and associative auditory cortices were not significant. Additional work is required to determine exactly why the category-specific differences arose. Nevertheless it is interesting that this apparent semantic difference was present in cortical regions associated with perceptual rather than semantic processing of sound (where more medial and anterior temporal activation might have been expected for living things; Gainotti, 2000; Devlin et al., 2002). This is at least suggestive that the effects stem from differences in the weighting of perceptual knowledge about the stimuli.

This current study also partially supports the findings of Lewis et al. (2007) in that the temporal lobe activity for animals was bilateral rather than unilateral (although activation was significantly stronger on the right side). While we also found no reliable activity for environmental sounds relative to animals, they reported greater left lateralised activity for environmental sounds. The two studies differed in the type of ‘non-living’ sounds stimuli used, as Lewis et al. employed only tool sounds while we used a mixture of household objects, vehicles and tools. The failure to find increased activation here, compared with animal sounds, may stem from the environmental sounds being more similar to animal sounds here (cf. Lewis et al., 2005). Alternatively it may be that there is a specific response to tools that was diluted here.

Auditory imagery

The imagery conditions resulted in no significant activation relative to baselines. This might be because the silent period was too short to allow for development of an auditory image or because participants recruited at least some overlapping processes in the baseline and imagery conditions. An account in terms of the timing of imaging does not seem very plausible. Previous behavioural studies indicate that participants take approximately two seconds to rate the vividness of their auditory imagery for sounds in the categories used in the current study. Also participants received a condition cue before each trial, to enable them to prepare to imagine the target as soon as it was presented. All participants reported they could do the task in the scanner, and gave a full range of vividness ratings, suggesting that participants could imagine the sounds. Therefore it is unlikely that participants were unable to imagine the target sound during this period.

The argument for similar processes in the baseline and imagery conditions seems more plausible. The baseline condition controlled for activation associated with button presses and word reading but participants may still have used inner speech, and this could mask any differential activity in the auditory imagery condition. To determine whether this occurred, this deactivation was plotted (i.e. by contrasting the baseline condition with the imagery condition). This revealed significant activation in the bilateral STG, which has been implicated in inner speech processes (Shergill et al., 2001). Future studies should control for inner speech whilst still matching the task demands of the auditory imagery task (e.g. by using a visual imagery task, such as mental rotation).

Despite this lack of main effect relative to baseline, animal sound imagery tended to evoke greater activation than imagery for environmental sounds within right auditory association regions, similar to that found in the perception conditions. This is consistent with imagery and perception using similar neural circuits. Bunzeck et al. (2005) also found overlap between imagery and perception for environmental sounds in their study. The authors reported that both imagery and perception of environmental sounds resulted in bilateral activation of the associative cortices, but that perception alone resulted in activity in the primary auditory cortices. The current study revealed strong bilateral activation in the primary auditory cortices during sound perception, but no significant activation in these regions during imagery, similarly to Bunzeck et al. (2005).

Of key interest in this study was whether vividness ratings were associated with changes in neural activation. Modulation of activation by imagery ratings (i.e. ratings analysis) investigated the regions responsive to higher ratings. The left MFG and the bilateral insula lobes showed the most pronounced activation. Previous studies of auditory imagery also report MFG activity (Yoo et al., 2001; Shergill et al., 2001). The MFG is also part of the dorsal lateral prefrontal cortex (DLPFC) and has been implicated in auditory working memory tasks (Arnott, Grady, Hevenor, Graham, & Alain, 2005; Grasby et al., 1994), so activation here may reflect general demands on working memory. Previous studies of inner speech also found bilateral insula lobe activity (Aleman et al., 2005; Shergill et al., 2001) and abnormal activation of this region has been implicated in auditory hallucination development (Nagai, Kishi, & Kato, 2007). A study of non-verbal auditory imagery also found insular activation, which linked the activation to sound retrieval processes (Hoshiyama, Gunji, & Kakigi,

2001) and the allocation of higher-level cognitive processes to auditory attention (Voisin, Bidet-Caulet, Bertrand, & Fonlupt, 2006). Therefore it appears that these areas may reflect the generation and maintenance of non-verbal auditory imagery.

The PPI analysis used the right auditory cortex as seed point, given that this area showed activation in both the perception vs. baseline contrast and in the animal sound imagery vs. environmental sound imagery contrast. This analysis revealed significant connectivity between the seed region and areas in the frontal lobe, when contrasting between imagery and perception conditions. This analysis indicates involvement of the frontal regions in either generation of auditory image or (possibly) in evaluating the strength of the images for the rating task. Enhanced connectivity between frontal and auditory cortex would be consistent with generation of stronger images for animals, through reciprocal links between these areas, or with stronger images in auditory cortex being fed-through to more frontal regions for the evaluation process.

In conclusion the current study supports and extends the findings of previous studies of auditory imagery and perception for non-verbal sounds. Previous studies have demonstrated distributed regions of activity to different sound categories and some overlap between imagery and perception. In such studies, acoustic properties of the sounds varied greatly between categories, therefore it is unclear whether dissociations between the sound categories resulted from acoustic or semantic differences between the stimuli. The current study balanced the acoustic properties across the stimuli but still demonstrated greater activation for animals than environmental sounds, and this tended to occur across both imagery and perception conditions. Individual differences in imagery vividness were related to altered activity in the MFG and insula cortex.

The study also showed increased connectivity between frontal and auditory cortex, during imagery conditions, suggesting a mediating role for these connections in either creating auditory images or in evaluating their strength.

Chapter Four Association between schizotypy and imagery vividness

The association between auditory imagery and hallucinations has long been hypothesised, but findings have been contentious, as some studies find strong associations and others find little link between the two. The current study makes comparisons between the sample of participants investigated in the current thesis and participants in previous studies in terms of schizotypy, hallucination proneness and anomalous experiences. Having established that the current sample of participants was similar to those in previous studies, the study then assessed the association between auditory imagery vividness and the factors mentioned above. The O-LIFE investigated schizotypy and positive traits of psychosis, the LSHS-R investigated hallucination proneness and the CAPS investigated anomalous experiences. In addition the VVIQ and a novel auditory imagery questionnaire assessed auditory and visual imagery vividness. The study showed a good association between the measures of schizotypy, hallucination and anomalous experiences. When correlated with imagery measures, people with greater positive symptom traits had a trend to higher auditory imagery vividness, and when grouped by positive symptoms, the high scoring group had significantly higher auditory imagery vividness than the low scoring group. However there was no association between either imagery measure and hallucination proneness. Assessment of the separate anomalous experiences from the CAPS revealed an association between auditory imagery vividness, sensory flooding and temporal lobe symptoms.

Experiment 5. Association between schizotypy and imagery vividness

Introduction

Hallucinatory experiences are often associated with psychiatric symptoms, such as bipolar depression, post traumatic stress disorder and schizophrenia (Beck et al., 2003). Indeed, hallucinations are a first-rank symptom for the diagnosis of schizophrenia (Schneider, 1959), and between 60 and 70% of such patients experience hallucinations (Bentall, Jackson, & Pilgrim, 1988; Sartoriu, Shapiro, Barrett, & Kimura, 1972). Hallucinations can occur in all modalities; however auditory hallucinations are most common, followed by visual, tactile and olfactory hallucinations (Mueser et al., 1990).

Auditory hallucinations can also occur in the absence of other psychiatric symptoms however. Romme and Escher (1989) found that many people consider hallucinations to be a normal part of everyday life, find them beneficial and do not have other experiences associated with psychiatric conditions. A number of surveys have been conducted to determine the prevalence of hallucinations in the normal populations. Such estimates vary between 10% and 25% of people in the normal population experience hallucinations (Tien, 1991; Barrett et al., 1992; Young et al., 1986). Therefore hallucinations could be considered to be a normal experience for many people.

Measurement of schizotypy, hallucination proneness and anomalous experiences

Like hallucinations, other schizophrenia symptoms are hypothesised to lie on a continuum with psychiatric symptoms, so that people can have schizotypal personality factors, which are at a subclinical level (Meehl, 1962).

A number of measures have been developed to assess schizotypal personality characteristics, such as the Perceptual Aberration Scale (Chapman et al., 1978) and the Schizotypal Personality Scale (Raine, 1991). The Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE) incorporated items from other measures of schizotypy, including the LSHS, STQ and the Perceptual Aberration Scale (Mason et al., 1995). The O-LIFE assesses key schizotypy traits related to positive and negative symptoms, and included four subscales, to assess Unusual Experiences, Cognitive Disorganisation, Impulsive Non-conformity and Introvertive Anhedonia. Initial investigations with this scale revealed good internal reliability and good association with other measures of psychosis, suggesting that the scale is a reliable and valid measure of schizotypy.

Other scales focus on hallucination-like experiences directly. The most commonly used measure of hallucination-like experiences is the Launay-Slade Hallucination Scale (Launay et al., 1981) . The original scale determined whether prison inmates have a higher incidence of hallucinatory experience than participants than in the normal population, but it has since been used to investigate hallucinatory experiences in a range of other participants and situations. For instance a later study by Bentall and Slade (1985) assessed hallucinatory experiences of 150 male non-psychiatric participants. Participants completed the LSHS on two occasions separated by up to three weeks, which revealed good correlation between the two versions of the questionnaire, suggesting that the LSHS measure is reliable and measures a stable trait. In addition scores on the LSHS correlated with other schizotypy measures, showing that it is a reliable measure.

This scale focuses primarily on the incidence of auditory hallucination experiences, despite the fact that hallucinations can be experience in any modality. In addition it does not take into account the frequency and distraction evoked by such experiences, which may give a better indication of the severity of such experience.

A more recent measure focuses on other types of anomalous experiences, as well as auditory hallucination and obtains a measure of frequency of experiences and distraction evoked. The Cardiff Anomalous Perceptions Scale (CAPS, Bell et al., 2006) assesses anomalous experiences in all sensory modalities, but does not directly ask about schizotypy or hallucination experiences. Participants give yes or no responses to the questions and where participants give a positive response, they also rated the experience using a 5-point rating scale, on three dimensions: the distress, intrusiveness and the frequency that the experience occurs. Bell et al. (2007) tested this scale with a group of non-psychiatric participants and a psychiatric group of participants and found that the total score on this scale correlated well with other measures of schizotypy. Interestingly in this study, the authors do not investigate the relationship between the different dimensions of the scale (i.e. total score, distress, intrusiveness and frequency ratings) and the other measures investigated.

Relationship with schizotypy and auditory imagery

All of the above scales appear to have a significant imagery component to them. The O-LIFE includes questions such as “Are your thoughts sometimes so strong that you can almost hear them?” in the UE scale, while the CAPS includes items such as “Do you ever hear noises or sounds when there is nothing about to explain them?” In addition both PCA’s of the LSHS seem to suggest a factor related to vivid imagery (i.e. "Vividness of daydreams", Levitan et al., 1996; Aleman et al., 2001).

Previous studies suggested vivid imagery has a key role in the development of hallucinatory experiences (Mintz and Alpert, 1972). Experimental investigation of the relationship between hallucination proneness and imagery abilities has produced mixed results however, with some studies finding a positive relationship between scores on the LSHS and imagery questions (Mintz & Alpert, 1972; Roman & Landis, 1945), others finding a negative relationship (Starker & Jolin, 1982) and some finding no relationship (Brett & Starker, 1977). One issue with these previous studies is that the LSHS focuses on with auditory unusual experiences, but few studies focus on their association with auditory imagery vividness, tending to incorporating more general measures of imagery, including all modalities of imagery (Aleman, Bocker, & de Haan, 1999; Aleman & de Haan, 2004).

The current study

The current study had two aims. Firstly the study determined whether the sample of participants matched that of previous studies in terms of hallucination proneness, anomalous experiences and schizotypy. This is in order to rule out that any association between these factors and imagery vividness is due to an idiosyncratic sample, rather due to true association. Secondly previous research of imagery abilities in association with schizotypy is lacking. Therefore the current study also looked at the associations between schizotypy, hallucination and imagery vividness, incorporating measures of schizotypy (O-LIFE) hallucination-like experiences (LSHS-R) and other anomalous experiences (CAPS).

Method

Participants

18 male and 76 female (18 – 34 years old) from the University of Birmingham completed both the CAPS questionnaire and the LSHS questionnaire as part of other experiments. 54 of these participants also completed the O-LIFE and 64 completed the imagery questionnaires.

Results

Part 1: Comparison to previous studies

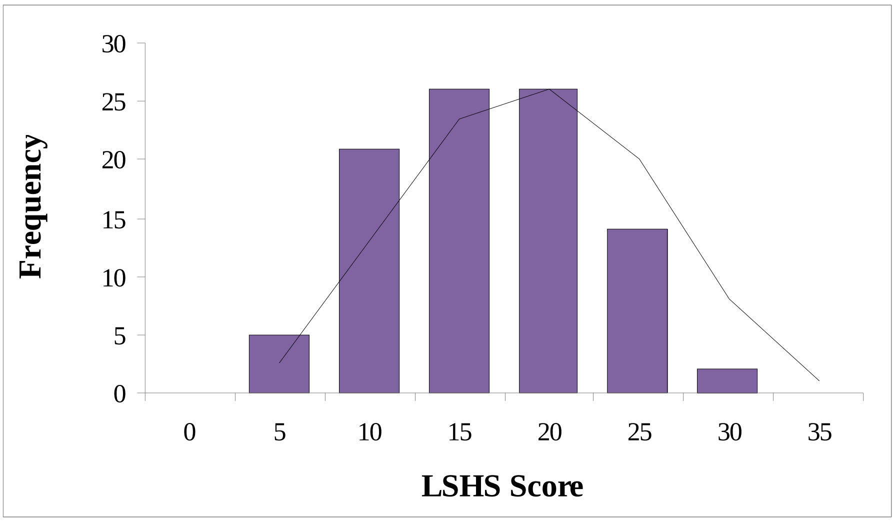


Figure 18: Distribution of LSHS scores in the current study

LSHS: 94 participants completed the LSHS, revealing a mean score of 14.6 (SD = 5.8). Figure 18 shows that the LSHS scores were normally distributed.

The population of this study most resembled that of Aleman et al. (2001) (i.e. undergraduate students). In Aleman et al.'s (2001) study 243 participants completed the LSHS, and the mean score was 13.9 (SD = 6.7). The mean, standard deviation and number of participants in Aleman et al.'s study were used to compare their results to those of the current study, using an independent samples t-test[5]. The scores of these two samples did not differ significantly $t(190)=0.957$, $p=0.341$.

O-LIFE:

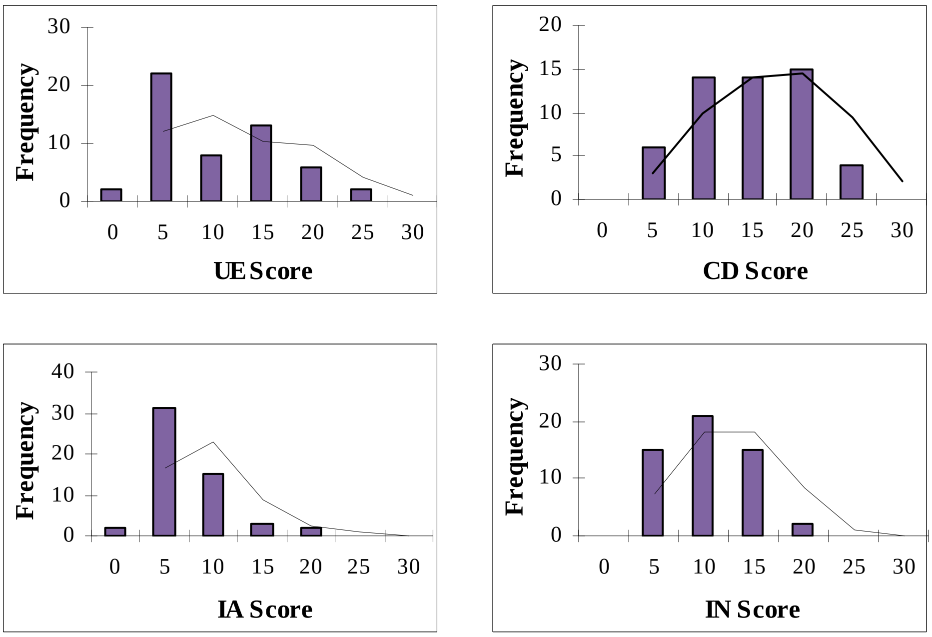


Figure 19: Distribution of O-LIFE subscale scores in the current study

[5] This analysis was conducted using an online t-test calculator, http://www.quantitativeskills.com/sisa/statistics/t-test.htm

54 participants completed the O-LIFE questionnaire. Figure 19 shows that the distributions for the CD and IN scales were normally distributed, whilst the UE and IA subscales were slightly skewed towards lower scores.

Table 11 shows the mean and standard deviation for each subscale for the current study and for Mason and Claridge's (2006) data. The mean, standard deviation and number of participants in Mason and Claridge's study were used to compare their results to those of the current study, using independent samples t-tests. Comparison between the two studies revealed no significant differences in scores for the UE and IN subscales (UE: *t(54)=0.118*, *p=0.907;* IN: *t(55)1.256*, *p=0.215*). Scores in the current study were significantly higher for the CD subscale however, *t(55)=2.518*, *p=0.015*, and significantly lower for the IA subscale, *t(56) = 2.632, p=0.011*.

Table 11: Comparison between Mason et al. (2006) and the current study (standard deviations in brackets)

Measure	**Current Study (N = 53)**	**Mason et al. (2006) (N =1926)**
UE	8.7 (6.1)	8.8 (6.2)
CD	12.7 (5.7)	10.7 (5.9)
IA	5.0 (3.8)	6.4 (4.5)
IN	8.4 (4.0)	7.7 (4.1)

CAPS:

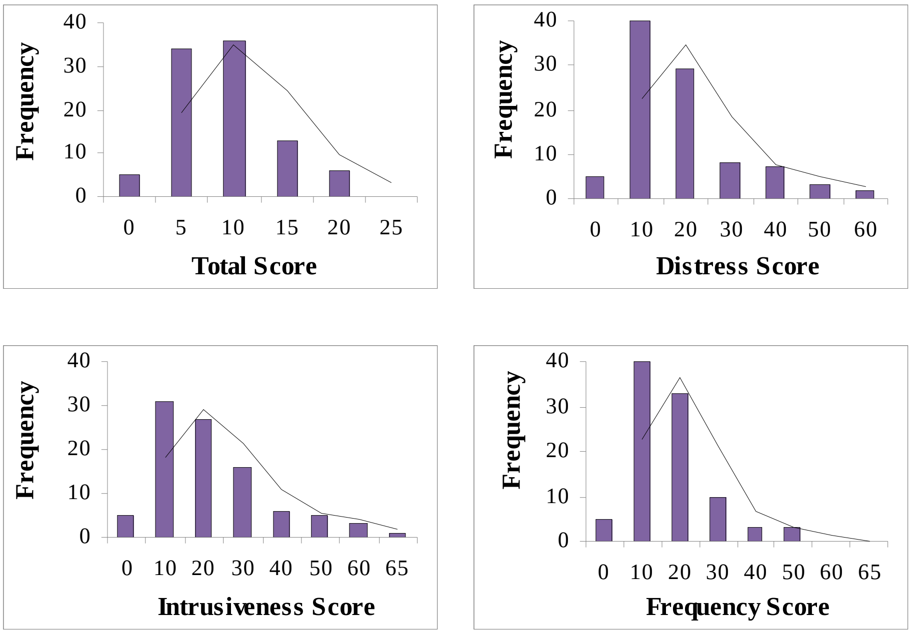


Figure 20: Distribution of CAPS subscale scores in the current study

94 participants completed the CAPS questionnaire. Figure 20 shows that all the CAPS subscales scores were slightly skewed towards the lower scores. Table 12 shows the mean and standard deviation for the current study and for Bell et al.'s (2006) study. Comparison between the two studies reveals no significant differences on any of the subscales *($p>0.05$)*.

Table 12: Comparison between Bell et al. (2006) and the current study (standard deviation in brackets)

Measure	Current Study (N = 94)	Bell et al. (2006) (N = 337)
Total	6.9 (4.7)	7.3 (5.8)
Intrusiveness	14.4 (12.2)	15.5 (14.5)
Distress	17.1 (14.4)	18.0 (17)
Frequency	12.8 (10.0)	14.6 (14.2)

Correlations between the schizotypy scales

The separate scales were then correlated with each other and compared to other studies that have also compared the scales. Table 13 shows a correlational matrix between the subscales of these measures. Due to the large number of comparisons, a more stringent criteria of p<0.01 was used to accept the significance of the correlations. This revealed good correlation between dimensions and subscales within each measure (i.e. CAPS, O-LIFE and AVIQ). In addition there were positive associations between the measure of hallucination proneness (LSHS) and the measures of anomalous experience (all CAPS dimensions) and schizotypy (all O-LIFE subscales, except IA). The CAPS dimensions also correlated well with the O-LIFE subscales (except the IA subscale).

Table 13: Correlational matrix for LSHS, CAPS subscales and O-LIFE subscale (r-values)

		LSHS	**CAPS**				**O-LIFE**			
Measure			**T.**	**D.**	**I.**	**F.**	**UE**	**CD**	**IA**	**IN**
LSHS (N = 94)		-								
CAPS (N = 94)	**T.**	0.29**	-							
	D.	0.33**	0.88**	-						
	I.	0.37**	0.91**	0.96**	-					
	F.	0.38**	0.91**	0.85**	0.91**	-				
O-LIFE (N= 54)	**UE**	0.59**	0.58**	0.61**	0.59**	0.57**	-			
	CD	0.52**	0.42**	0.44**	0.43**	0.40**	0.72**	-		
	IA	0.04	0.08	0.17	0.13	0.06	0.29*	0.30*	-	
	IN	0.48**	0.43**	0.43**	0.44**	0.43**	0.62**	0.54**	0.17	-

T. = Total; D. = Distress; I. = Intrusiveness; F. = Frequency
*** - p<0.05, ** - p<0.01**

Bell et al. (2006) also investigated the correlations between the CAPS total score, LSHS and O-LIFE subscales. Table 14 shows these correlations, as well as those of the current study.

Table 14: Comparison of correlations with CAPS total score between Bell et al. (2006) and the current study

CAPS total score		LSHS	O-LIFE			
			UE	**CD**	**IA**	**IN**
Current study	**Pearson's r**	0.29**	0.58**	0.42**	0.08	0.43**
	N	94	53	53	53	53
Bell et al. (2007)	**Pearson's r**	0.65*	0.57*	0.36*	0.03	0.20**
	N	288	170	170	169	171

These comparisons reveal that, as in the current study, correlations between the CAPS total score and the LSHS, and the O-LIFE subscales of UE, CD and IN were significant.

Further analysis then determined whether these correlations differed significantly from Bell et al.'s (2007) study. This revealed a borderline significant difference between Bell et al.'s (2007) study and the current study, in the correlations for the CAPS total score and the IN score, $z=1.596$, $p =0.055$. The correlation was stronger in the current study in comparison to Bell et al.'s (2007) study.

Part 2. Association between hallucination proneness, schizotypy and imagery abilities

69 participants completed the imagery questionnaires. The mean score on the auditory imagery subscale was 78.7 (SD=13.2) and the mean score on the visual imagery subscale was 85.2 (SD=12.1).

Table 15: Correlations between the measures of hallucination proneness and schizotypy and imagery measures

Measure		AI	VI
LSHS (N=69)		0.15	0.11
CAPS (N=69)	**Total**	0.18	0.07
	Distress	0.11	0.03
	Intrusiveness	0.14	0.02
	Freq.	0.10	-0.04
O-LIFE (N=45)	**UE**	0.28#	0.10
	CD	0.18	0.16
	IA	0.08	0.07
	IN	0.07	-0.20

Table 15 shows the correlations between the hallucination and schizotypy measures and the imagery subscales. No correlations met the significance criteria of $p<0.01$, however one correlation approached significance. This showed an association between higher scores on the UE subscale of the O-LIFE and higher auditory imagery scores, *r(45)=0.280, p=0.068*.

Table 16: Correlations between the imagery subscales and the CAPS subscale categories(r-values; N=69, # - p<0.07, * - p<0.05, ** - p<0.01)

Measure		CAPS category 1	2	3	4	5	6	7	8	9
AI	**T.**	0.15	0.06	0.16	0.06	0.11	-0.01	0.24#	0.02	0.22#
	D.	0.12	0.02	0.18	0.04	0.03	-0.72	0.33**	-0.13	0.14
	I.	0.16	0.05	0.17	0.01	0.08	-0.08	0.29*	-0.03	0.23#
	F.	0.09	0.05	0.16	0.06	0.00	-0.12	0.20	-0.01	0.11
VI	**T.**	0.08	0.01	0.03	0.03	0.04	0.08	0.15	-0.02	0.04
	D.	0.10	-0.02	0.14	0.01	0.01	-0.01	0.25*	-0.19	-0.03
	I.	0.11	-0.06	0.10	-0.03	-0.02	-0.03	0.18	-0.10	0.04
	F.	0.04	-0.12	0.08	-0.03	-0.07	-0.04	0.04	-0.04	-0.06

1.Changes in levels of sensory intensity; 2. Having a non-shared sensory experiences; 3. Inherently unusual or distorted sensory experiences; 4. Sensory experience from an unexplained Source; 5. Distortion of form of own body and of external world; 6. Verbal hallucinations; 7. Sensory flooding; 8. Thought echo and hearing thoughts out loud; 9.Temporal lobe disorder

Table 16 shows the correlations between the four subscales on the CAPS categories and the auditory and visual imagery scores. Again the criteria of $p<0.01$ was used to accept significance. Only one comparison met this criterion. This showed an association between higher auditory imagery scores and higher distraction scores on the “Sensory Flooding” (7) category. Four other comparisons approached significance. These were between higher auditory imagery scores and the total score on the “Sensory Flooding” category, and between higher auditory imagery scores and increased total score and intrusiveness score for “Temporal Lobe” (9) category. A trend was also found between higher visual imagery scores and higher scores on the distress ratings for the “Sensory Flooding” (7) category.

Multiple Regression analysis

The following analysis determined how the different measures of schizotypy were associated with imagery vividness. Analyses were conducted separately for Auditory Imagery scores and Visual Imagery scores

Auditory Imagery

This analyses used auditory imagery score as the dependant variable, and LSHS score, the O-LIFE subscale scores (UE, CD, IA and IN) and the CAPS total score[6] as the independent variables. These variables accounted for only 25% of the variance in auditory imagery scores, $F(6, 44) =0.929, p = 0.485$. No regression coefficients were significant.

Visual Imagery

Again, visual imagery score was entered as the dependant variable, and LSHS score, the O-LIFE subscale scores (UE, CD, IA and IN) and CAPS scores (total score, Intrusiveness score, distress score and frequency score) as the independent variables.

[6] Because of the high intercorrelations between subscales of the CAPS, tolerance level was very low for these measures, therefore only the CAPS total score was included in the model. For all other measures, VIF and the tolerance were within acceptable levels.

These variables accounted for 40.1% of the visual imagery variance, $F(6, 44) = 3.018$, $p = 0.016$. Two regression coefficients reached significance: LSHS score, $\beta = 0.408$, $t = 2.441$, $p=0.031$, the IN subscale of the O-LIFE, $\beta = -0.529$, $t=-2.809$, $p=0.008$.

Summary

Table 17 summarizes the findings of each part of the study.

Table 17: Summary of results

Part	Findings
Part one: Comparison to previous studies	Good association between current study and previous studies using LSHS-R and CAPS, and the UE and CD subscales of the O-LIFE
	Sample in the current study had significantly higher CD scores and lower IA scores on the O-LIFE subscale, compared to Mason and Claridge's (2006) sample
	Good correlation between LSHS-R, subscales of CAPS and all subscales of O-LIFE (except IA)
Part two: Association between schizotypy and imagery vividness	Near significant positive correlation between AI and UE
	Significant correlations between AI, the sensory flooding category and the temporal lobe category on the CAPS
	Significant correlation between VI and sensory flooding distress score
	No significant regressors on auditory imagery scores
	Significant regressors of LSHS score and the IN subscale (from O-LIFE)

Discussion

This study had two aims: (i) to compare participants from the current studies to previous investigations of schizotypy and hallucination proneness and (ii) to assess the association between imagery abilities and these measures. The analyses revealed similarities in the data from participants in this study and those of previous studies. The majority of mean values did not differ across studies, all measures correlated well with each other and the correlations between measures did not differ across studies. The study revealed two differences however. Here, compared with Mason and Claridge (2006) scores were significantly lower on the IA (i.e. negative symptoms) scale and higher on the CD scale, though the scores most related to hallucination experiences (i.e. the LSHS-R and the UE scale of the O-LIFE) did not differ. Thus the present participants were similar to prior studies in terms of hallucination proneness. The implications of the differences in IA and CD are unclear. Possibly however, the combination of decreased negative symptoms and increased cognitive disorganisation may increase the likelihood of hallucinatory experiences, as this combination of traits may lead to confusion between imagined events and real events. In addition CD correlates highly with hallucination proneness, suggesting an element of cognitive disorganisation in the experience of hallucinations.

Association with mental imagery

Correlations between auditory and visual imagery and hallucination proneness revealed no significant differences between high and low scorers. Sack, van de Ven, Etschenberg, Schatz, and Linden (2005) also found no correlation between self-rated imagery vividness (assessed using Betts' QMI) and hallucination proneness (assessed using the LSHS-R). Sack et al. did find that hallucinating patients had more vivid imagery in all modalities (including auditory imagery) than control participants, however. Aleman, Nieuwenstein, Bocker, and de Haan (2000) also investigated the association between imagery and hallucination proneness, using the visual and auditory subscales of the QMI and the VVIQ. They found that high hallucination prone participants had significantly more vivid visual imagery on the VVIQ, but did not differ in their scores on either the visual or the auditory subscale of the QMI. The association between the auditory imagery subscale and hallucination proneness in the current study therefore seems to support the findings of the previous studies associations between the auditory subscale of the QMI and the LSHS-R.

The correlations conducted on the auditory and visual imagery vividness scales and the O-LIFE subscales revealed a borderline significant relationship between auditory imagery and the UE subscale only. This indicated an association between increased auditory imagery abilities and increased reports of unusual experiences. Further, multiple regression analysis showed little association between auditory imagery and different measures of schizotypy, as there were no significant regressors. For visual imagery however, LSHS-R score and the IN subscale of the O-LIFE questionnaire emerged as significant regressors. This suggests that visual imagery may be more related to hallucination proneness than auditory imagery.

The O-LIFE has not been used before to assess auditory imagery abilities associated with schizotypy. Van de Ven and Merckelbach (2003) investigated the association between schizotypy, hallucination, fantasy proneness and mental imagery in hallucinatory reports of undergraduate students. Mental imagery vividness (measured using the QMI) correlated with both hallucination proneness (measured using the LSHS-R) and schizotypy (measured using the STQ, Claridge et al., 1984). Though the STQ does not specifically differentiate between positive and negative traits of schizotypy, Van de Ven et al.’s (2003) study offers some support for the association between mental imagery vividness and positive schizotypy, as in their study, the correlation between imagery and hallucination proneness was also significant.

An interesting finding of the present study was the correlation found between the CAPS categories and the auditory imagery questionnaire score. Auditory imagery correlated with all subscales of the ‘sensory flooding’ category. Therefore the study suggested an association between increased auditory imagery vividness and increased stimulation of sensory experiences. In addition, a previous study suggested the involvement of sensory flooding in hallucination development , as sensory overload may lead to misperceptions of sensory stimuli, leading in turn to the experience of hallucination (Freedman et al., 2002). Thus this finding in the current study offers support for the theory that vivid auditory imagery and hallucinations exist on extreme ends of a continuum (Bentall et al., 1985).

Limitations

One possible reason for the lack of strong association between imagery and schizotypy is the subjective nature of the imagery questionnaire. For example, people who are higher in schizotypy may experience more vivid imagery than those low in schizotypy, but have equal subjective impressions of how vivid their imagery is. This would explain previous studies which found differences in performance on objective imagery tasks, but no difference in subjective auditory imagery abilities.

Another possibility is that it is not vividness of auditory imagery per se that influences hallucinatory experiences, but rather how those images are attributed. Mintz and Alpert (1972) suggested that vivid imagery, in combination with reality monitoring errors, leads to hallucinatory experiences. Further research is needed to investigate the nature of the association between reality monitoring errors and imagery vividness.

Conclusion

In conclusion, the current study established that the sample of participants assessed in this thesis is generally representative of samples assessed previously. In particular the current sample resembled previous samples in terms of positive schizotypy, hallucination proneness and anomalous experiences, though the current participants had lower levels of negative schizotypy and higher levels of cognitive disorganisation than previous studies. The current study also established the association between the schizotypy measures and imagery vividness. Regression analyses showed an association between auditory imagery and frequency of anomalous experience and, as significant regressors of hallucination proneness, impulsive non-conformity and frequency of anomalous experiences on visual imagery. This analysis suggests that visual imagery may be more associated than auditory imagery with hallucinatory experiences. In addition there was an association between auditory imagery and sensory flooding, suggesting and association between that increased auditory imagery vividness and over-stimulation from sensory experiences.

Chapter Five. Sound detection, vividness and hallucination proneness

Abstract

Previous studies have suggested that hallucination-prone people have a greater bias towards thinking sounds are present in noise, compared to those who do not. The following study employed a sound detection paradigm, requiring high and low hallucination prone participants to detect threshold level sounds in white noise. The sounds differed in auditory imagery vividness and familiarity and participants either received a cue to what the target was, or no cue. Vividness and familiarity affected response bias but only familiarity affected sensitivity. The presence of a cue to the sound increased both response bias and the confidence of responses. This suggests that cognitive factors associated with sounds influence participants belief in the presence of sound in noise, leading to a greater bias in conditions where sounds are easier to imagine. This is a general trait however, which is unrelated to hallucination proneness.

Experiment 6. Sound detection, vividness and hallucination proneness

Introduction

The data presented in Chapter 2 demonstrated the effects of sound category, familiarity, item cueing and the perception of other sounds on rated auditory imagery vividness (i.e. the subjective clarity of an auditory image, Aleman et al., 1999). Overall music and speech images received higher vividness ratings compared to animal and environmental sounds, familiar images received higher vividness ratings than unfamiliar images and picture cues to sounds resulted in higher imagery vividness compared to names cues. Vividness ratings however rely on subjective judgments and do not necessarily reflect perceptual sensitivity to sounds. Hence rating studies do not necessarily indicate that imagery for sound interacts with the perceptual processing of sounds. To assess whether there are effects on perceptual processing, Experiment 3b examined how imagery vividness and familiarity affect the detection of sounds in noise. The study revealed no difference in the bias to reporting the presence of familiar low and high vividness sounds, however the bias was greater for unfamiliar high vividness sounds compared to low vividness sounds. In addition, participants were more confident in having heard the sounds in cue conditions. These factors did not affect the sensitivity to the sound however. In the following study the same sound detection task assessed how proneness to hallucination-like experiences interacted with vividness, familiarity and detection cues.

Proneness to auditory hallucinations

Auditory hallucinations are perceptual experiences of sounds in the absence of an external stimulus, and are a core symptom in the diagnosis of psychotic disorders such as schizophrenia. Indeed, 65% of schizophrenic patients experience auditory hallucinations at some time during their lives (Slade & Bentall, 1988). Such hallucinations can range from vague perceptions of noises (such as buzzing and banging sounds) to more complex sounds such as music or speech (Beck et al., 2003). Speech is the most commonly reported auditory hallucination and can range from single words to full sentences (Beck et al., 2003).

Previous studies suggest that vivid imagery is more perceptual-like in quality than less vivid imagery, so that vivid images are more likely to be confused for actual percepts, which in schizophrenic participants, may lead to hallucinations (Aleman et al., 2003). The precise link between imagery and hallucination is yet to be determined however. Some studies suggest that patients who suffer from hallucinations have a more vivid imagery system than normal people (Bocker, Hijman, Kahn, & de Haan, 2000; Mintz et al., 1972), whereas others suggest that the imagery system of such individuals is weaker than those who do not experience hallucinations (Starker et al., 1982; Brett et al., 1977), so that when experience of mental imagery occurs it is so unusually vivid that it is perceived as coming from an external source.

Many studies focus on whether perception and imagery are more similar to each other in hallucinating compared to non-hallucinating participants. One way to do this is to investigate the effect of imagining a sound on the ability to detect it in noise. Previous studies of normal participants found that imagining a target either inhibits (Segal et al., 1970) or facilitates (Farah et al., 1983) detection of sounds in noise, depending on the experimental conditions, showing that imagery for sounds can interact with perception, and influence detection of the target item.

Studies have also investigated the differences between high and low hallucination prone participants in their ability to detect a sound in noise, to determine whether the interaction between imagery and perception differs in these two groups. The term 'reality monitoring' refers to the ability to distinguish real events from imagined events, and previous studies have linked this ability to the development of hallucinations in schizophrenic patients, and to hallucination-like experiences in control participants (Johnson et al., 1981). Mintz and Alpert's (1972) study suggested that vivid imagery in association with reality monitoring impairments combine to produce hallucinations in people with psychosis symptoms. Bentall and Slade (1985) also found reality monitoring deficits in both hallucinating psychosis patients and hallucination-prone normal participants. In this case, groups had a stronger bias to reporting the presence of an auditory word in noise, i.e. they were more likely to believe that imagined targets were real. Bentall and Slade rejected the theory that hallucinating patients have more vivid imagery however. They did not find a difference between hallucinating (and hallucination prone) and non-hallucinating participants in sensitivity to auditory targets, which they claimed would indicate that auditory images modulated sound detection.

Barkus et al. (2007) corroborated the findings of Bentall and Slade (1985), finding increased bias for normal participants with high hallucination proneness (scoring high on the LSHS, Launay et al., 1981; and the UE subscale of the O-LIFE, Mason et al., 1995) but no effects on sensitivity. Eight high hallucination prone participants also completed an fMRI version of this study which revealed activation in areas associated with auditory hallucinations and verbal imagery (i.e. middle temporal gyrus and the fusiform gyrus during false positive responses). Other evidence for effects on bias rather than perceptual sensitivity comes from the work of Merckelbach and van de Ven's (2001). These authors required participants to detect a song in noise and found that high hallucination prone individuals were more confident that they heard the song than low hallucination prone participants.

These studies show that hallucinating schizophrenic participants and non-schizophrenic participants who have hallucination-like experiences are more likely to believe they hear a sound in noise than non-hallucinating participants. Prior results suggest an effect on response bias more than on sensitivity, though the generality of this result requires further exploration. Very little research has focussed on categories other than single words and music. The present study measured signal detection across a number of object classes using individuals classed either as high or low in hallucination proneness. In addition, to contrast performance in individuals differing in hallucination proneness, we also assessed detection in relation to the vividness and familiarity of the sounds presented, and provision of a cue before the target provision. The effects of auditory vividness, familiarity and the cue can be attributed to the top-down influence of pre-knowledge (from long-term memory or from the cue) on perception and response bias. By assessing the effects of hallucination proneness in relation to these factors, we can determine whether the effects of hallucination proneness arise through top-down knowledge in a manner similar to these other variables.

Method

Participants

243 participants from the University of Birmingham completed the O-LIFE questionnaire (Mason et al., 1995). The scores on the Unusual Experience subscale were used to group participants, as this subscale correlates strongly with hallucination proneness (Bell et al., 2006). The experimenter then contacted participants who scored one standard deviation above and one standard deviation below the mean and asked them to participate in the experiment. 67 people participated in the full experiment (59 females, 8 males; mean age: 19.6). These participants were recruited from the University of Birmingham, in exchange for participation credit or money. All stated they had normal hearing in both ears. 38 of these participants also participated in the previous sound detection experiment.

Questionnaires

The participants completed three questionnaires, to assess different behavioural factors that may affect performance on the sound detection task.

1. The Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE) (Mason et al., 1995) assessed schizotypal personality traits, and includes four subscales: Unusual Experiences (UE; i.e. the positive symptoms of schizotypy), Cognitive Disorganisation (CD), Introvertive Anhedonia (IA; i.e. negative symptoms of schizotypy) and Impulsive Non-Conformity (IN). The study analysed the total score and the UE score only as these components are most related to hallucination-like experiences.

2. The Launay-Slade Hallucination Scale – Revised (Bentall and Slade, 1985). This 12-item scale measures proneness towards hallucination-like experience and previous studies link high scores on this scale with an increased bias towards responding positively in auditory signal detection tasks (Bentall et al., 1985).

3. An imagery questionnaire. This contained an auditory imagery section and a visual imagery section from the Vividness of Visual Imagery Questionnaire (VVIQ, Marks, 1973). Participants rated the intensity of their imagery for different items.

Signal Detection Task

The stimuli and design were the same as in Chapter 2.

Procedure

Chapter 2 contains a description of the full procedure of the sound detection task, however a brief description follows. The experimenter informed the participants were that they were participating in a sound detection experiment where they would detect familiar and unfamiliar sounds in white noise. The experimenter informed the participants that they may or may not hear the sound in the white noise and that if the sound was present it would either be at an audible level or a threshold level. For this reason, the experimenter told the participant that they should listen very carefully to the white noise. The experimenter informed participants that they would either receive with a valid cue of the target sound and picture (cue condition) or they would receive no cue (no cue condition). Participants received a prompt at the beginning of each trial, to inform them whether it was a cue or no cue condition. In the cued condition, the experimenter asked the participant to listen carefully for the target sound whilst listening to the white noise (see Appendix D for instructions given to the participants).

Once participants indicated whether the target was present or absent in the white noise, they then rated how confident they were in their answers, from one to three (three being confident and one being not at all confident). In the 'no cue' condition, the participants also indicated what the sound was and whether it was familiar to them. The participants also completed the two remaining questionnaires. Half completed the imagery questionnaire at the beginning of the first session, and half before the second session. The participants completed the LSHS-R at the end of the experiment to avoid the questionnaire affecting participant's behaviour towards the task, as it clearly asked about hallucination-like experiences. Participants were then grouped according to whether they were high hallucination prone, if they scored above the median score for the whole group on the LSHS-R or low hallucination prone if they scored below the median on the LSHS-R and participants were tested until equal numbers of participants were tested for the different sound combinations (i.e. AM, EV, AV, EM).

Results

67 people participated in the full experiment, but five were excluded from the analysis because of low hit rates (i.e. more than two standard deviations below the mean hit rate for items in the cue condition), two were excluded because they had negative d′ value, which suggests that participants are not performing the task properly, and four participants were excluded to ensure the sound combination groups were even (in

these cases, the participants with the highest or lowest LSHS-R score were excluded depending on the hallucination proneness group they were in).

Therefore, data from 56 people (52 female and 4 males, mean age 19.54 years old) was analysed in full.

Questionnaires

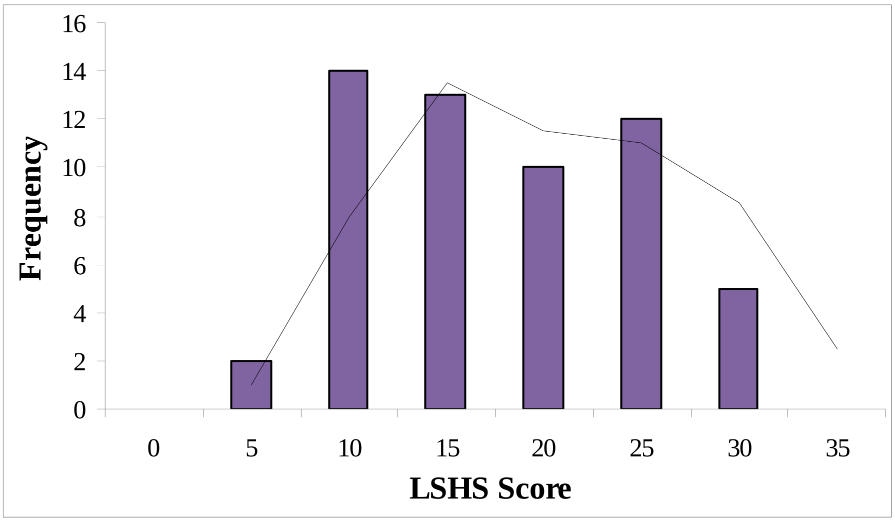


Figure 21: Distribution of LSHS scores

Participants were grouped according to their score on the LSHS-R. Figure 21 shows that the LSHS scores were normally distributed.

This resulted in 28 participants scoring above the median (> 15) on the LSHS-R (high hallucination prone) and 28 scoring on or below the median (< 15) on this scale (low hallucination prone). Participants that scored on the median were further grouped according to whether they scored above or below the median on the Unusual Experiences subscale of the O-LIFE questionnaire.

Figure 22 and Figure 23 show the distribution of the O-LIFE scores and the imagery questionnaire scores. These distributions appear normal, though the distribution for the IA score appears slightly skewed towards lower scores.

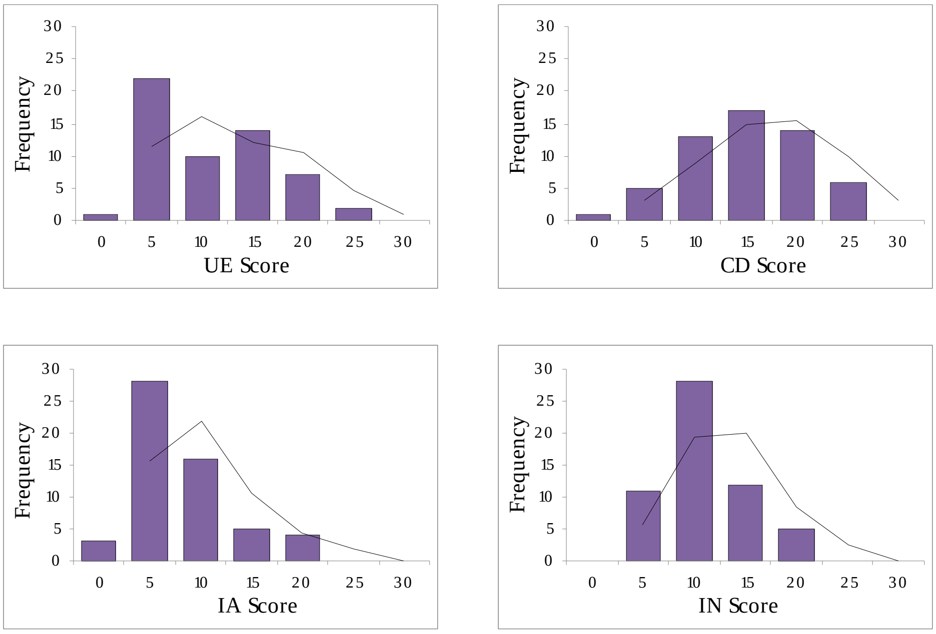


Figure 22: Distribution of O-LIFE subscale scores

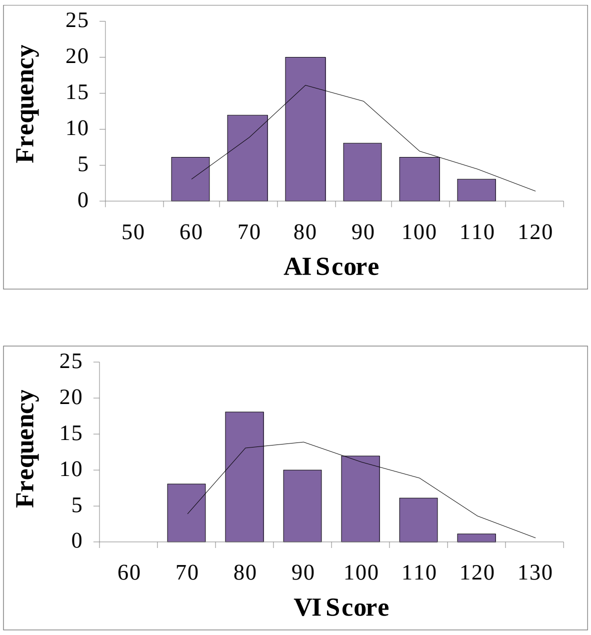


Figure 23: Distribution of imagery questionnaire scores
(AI =Auditory Imagery ; VI = Visual Imagery)

Table 18 shows the mean and standard deviations of the schizotypy, hallucination proneness and imagery questionnaires for each hallucination proneness group. These scores indicate that the high hallucination prone participants scored higher than the low hallucination prone participants on all the measures.

Table 18: Average scores on imagery questionnaire and O-LIFE (standard deviations in brackets)

		High H.P.	**Low H.P.**	**Total**
	LSHS	21.61 (3.91)	9.93 (3.25)	15.87 (6.89)
O-LIFE	**UE**	12.64 (5.68)	4.41 (4.24)	8.6 (6.49)
	CD	15.89 (5.41)	9.74 (5.11)	12.87 (6.07)
	IA	6.32 (5.49)	5.37 (3.40)	5.85 (4.57)
	IN	10.96 (3.78)	6.70 (3.15)	8.87 (4.06)
Imagery Questionnaires	**AI**	77.89 (14.08)	73.26 (11.42)	75.62 (12.94)
	VI	85.36 (15.17)	82.19 (10.97)	83.8 (13.25)

Table 19 shows the correlations between the schizotypy measure (O-LIFE), the positive symptoms subscale of the measure (UE), hallucination proneness (LSHS-R) and the imagery vividness measures (VI and AI).

Table 19: Correlations between schizotypy, hallucination proneness and imagery questionnaires (r-values)

	O-LIFE (N=56)				**Imagery Questionnaires (N=55)**	
	UE	**CD**	**IA**	**IN**	**AI**	**VI**
LSHS	0.665**	0.510**	0.072	.444**	0.214	0.228
UE	-	0.597**	0.264*	.603**	0.050	0.087
CD		-	0.409*	.474**	0.049	0.078
IA			-	0.038	-0.117	-0.240
IN				-	-0.086	-0.120
AI					-	0.645**

***= p<0.05;** = p< 0.001**

There were reliable correlations between the schizotypy measures and the hallucination proneness measure, and between the two imagery subscales. There were no correlations between either of the imagery subscales and the measures of schizotypy or hallucination proneness, however.

Sound Detection Task

The sound detection task was analysed in two ways[7]. The first analyses focused on the signal detection measures of sensitivity (d') and response bias (c) and the measure of confidence responses. These analyses determined how characteristics of the target sounds (vividness and familiarity), characteristics of the participants (hallucination proneness) and provision of detection cues influence the detection of target sounds in noise.

The second analyses focused on correlations to determine how reality testing differs between high and low hallucination prone participants, by correlating hit rate and confidence ratings. In addition further correlations, determined whether direct associations existed between the sound detection measures and vividness and familiarity ratings for each target item.

Like in Chapter 2, the analysis of this task began with the extraction of the overall hit and false positive rate for the high and low hallucination prone groups, in each condition. Figure 24 shows a plot of the hit rate against the false positive rate. This shows generally higher hit and false alarm rates for high vividness, unfamiliar sounds,

[7] Appendix F contains further analyses of the acoustic characteristic of target sounds and how these characteristics influence sound detection of high and low hallucination prone participants.

and little difference in these rates between high and low hallucination prone participants or between cueing conditions.

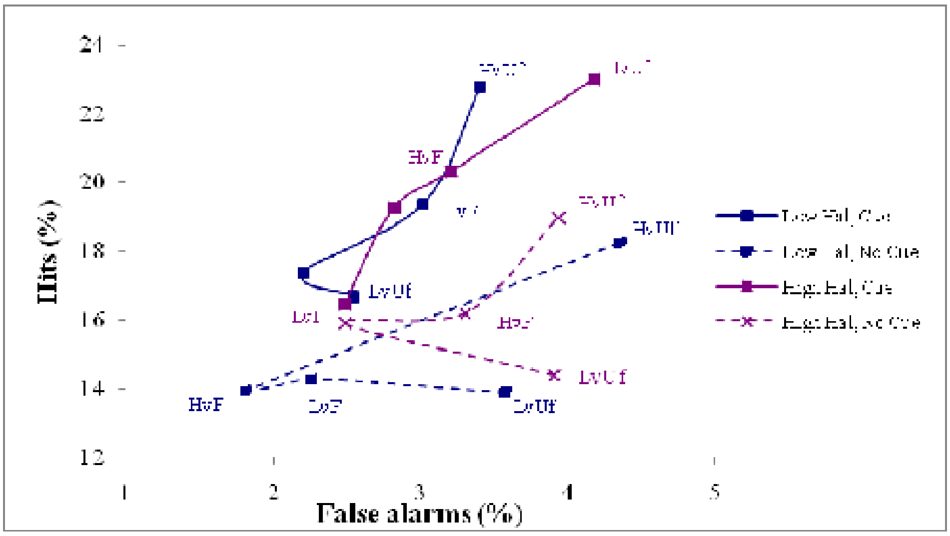


Figure 24: Plot of hits against false alarms, for each condition in sound detection task

Analysis one: The effect of vividness, familiarity and cue condition on sound detection of high and low hallucination prone participants

The hits and false alarm rates were used to calculate the c and d' measures, described in Chapter 2, using the method noted by Stanislaw et al. (1999). These analyses determined whether high and low hallucination prone participants differ in their sound detection performance, and how imagery vividness, familiarity and detection cues affected this performance.

Criterion

The criterion assesses the participant’s bias towards responding with a positive answer. A lower criterion indicates a greater the bias towards responding positively in an ambiguous signal plus noise environment and towards making more false positive responses.

Table 20: Criterion values in each condition, for high and low hallucination prone participants

		Low Imagery		High Imagery	
	H.P. Group	**Familiar**	**Unfamiliar**	**Familiar**	**Unfamiliar**
Cue	High H.P.	2.31 (0.78)	2.44 (0.80)	2.00 (0.49)	2.00 (0.52)
	Low H.P.	2.27 (0.72)	2.29 (0.80)	2.32 (0.89)	2.16 (0.76)
No Cue	High H.P.	2.44 (0.91)	2.19 (0.89)	2.40 (1.01)	2.11 (0.74)
	Low H.P.	2.50 (0.84)	2.38 (0.96)	2.66 (0.91)	2.03 (0.79)

Table 20 shows the average criterion for each hallucination proneness group, in each condition. For low vividness sounds (i.e. animal sounds and environmental sounds) high and low hallucination prone participants showed similar levels of performance. For high vividness sounds (i.e. music and speech sounds), low hallucination prone participants appeared to have a lower criterion than high hallucination prone participants (except when the sound was unfamiliar in the no cue condition).

Analysis of the criterion data consisted of a four-way split plot ANOVA. The within participants factors were vividness, familiarity and condition and the between participants factor was hallucination proneness. This analysis revealed a significant main effect of vividness, $F(1, 54) = 7.463, p = 0.008, \text{partial } \eta^2 = 0.121$, showing a greater bias to high vividness sounds. It also revealed a significant main effect of familiarity, $F(1, 54) = 10.020, p = 0.003, \text{partial } \eta^2 = 0.157$, showing that participants had a stronger bias towards stating that unfamiliar relative to familiar sounds were present. The main effect of condition was also significant, $F(1, 54) = 17.792, p < 0.001, \text{partial } \eta^2 = 0.248$; there was significantly larger bias towards thinking the sound was present, in the cue condition. There was no main effect of hallucination proneness group, $F(1, 54) = 2.088, p = 0.154, \text{partial } \eta^2 = 0.037$.

The analysis also revealed two significant interactions: between vividness and familiarity, $F(1, 54) = 13.204$, $p < 0.001$, *partial* $\eta^2 = 0.196$, and between familiarity and condition, $F(1, 54) = 7.656$, $p = 0.008$, *partial* $\eta^2 = 0.124$.

The analyses which follow decompose these interactions. For familiar sounds there was no significant bias difference between low and high vividness sounds, $t(55) = 0.395$, $p = 0.694$; however for unfamiliar sounds, there was a significantly greater bias towards high compared to low sounds, $t(55)= 4.226$, $p < 0.001$.

Analysis of the cue condition revealed that the bias did not differ between familiar and unfamiliar sounds $t(55) = 0.523$, $p=0.603$. In the no cue condition however the bias was greater for unfamiliar sounds, $t(55) = 3.758$, $p<0.001$. No other main effects or interactions were significant.

d′

D′ is a measure of sensitivity to a signal. A high d′ prime indicates a higher sensitivity to the target sound.

Table 21: Average d' in each condition, for high and low hallucination prone participants

		Low Imagery		**High Imagery**	
	H.P. Group	**Familiar**	**Unfamiliar**	**Familiar**	**Unfamiliar**
Cue	High H.P.	1.42 (0.80)	1.45 (0.74)	1.23 (0.46)	1.26 (0.54)
	Low H.P.	1.33 (0.70)	1.31 (0.78)	1.43 (0.82)	1.41 (0.73)
No Cue	High H.P.	1.41 (0.93)	1.09 (0.83)	1.40 (0.97)	1.20 (0.73)
	Low H.P.	1.43 (0.80)	1.28 (0.86)	1.53 (0.89)	1.11 (0.73)

Table 21 shows the average d' for the high and low hallucination prone group, in each condition. These scores suggest that the two groups did not differ in d' nor did vividness, familiarity or detection cues influence performance.

The d′ data were analysed using a four-way split plot ANOVA. The within participants factors were vividness, familiarity and condition and the between participants factor was hallucination proneness group. Firstly there was no significant main effect of vividness, $F(1, 54) = 0.076, p = 0.784$, partial $\eta^2 = 0.001$, condition, $F(1, 54) = 0.056, p = 0.814$, partial $\eta^2 = 0.001$, or hallucination proneness, $F(1, 54) = 1.165, p = 0.285$, partial $\eta^2 = 0.021$. There was a significant main effect of familiarity, $F(1, 54) = 4.0402, p = 0.049$, partial $\eta^2 = 0.070$. Participants were better at detecting familiar sounds. There was just one interaction, between familiarity and condition, $F(1, 54) = 5.168, p = 0.027$, partial $\eta^2 = 0.087$. In the cue condition there was no significant difference in sensitivity between familiar and unfamiliar sounds, $t(55) = 0.024, p=0.981$. The no cue condition however revealed that participants were significantly more sensitive to familiar compared to unfamiliar sounds, $t(55) = 2.726, p=0.009$. No other main effects or interactions were significant.

Confidence Ratings

Table 22: Average confidence ratings in each condition for the high and low hallucination prone groups (standard deviations in brackets)

		Low Imagery		High Imagery	
	H.P. Group	**Familiar**	**Unfamiliar**	**Familiar**	**Unfamiliar**
Cue	High H.P.	2.71 (0.33)	2.71 (0.34)	2.79 (0.37)	2.83 (0.33)
	Low H.P.	2.78 (0.14)	2.71 (0.22)	2.75 (0.21)	2.84 (0.15)
No Cue	High H.P.	2.57 (0.34)	2.51 (0.30)	2.55 (0.36)	2.66 (0.39)
	Low H.P.	2.64 (0.20)	2.57 (0.26)	2.45 (0.35)	2.66 (0.31)

A four-way split plot ANOVA was used to analysis the confidence ratings for hit responses. The within-participants factors were imagery vividness, familiarity, condition and the between participants factor was hallucination proneness group. There was only one reliable main effect, that of condition, $F(1, 54) = 70.803, p < 0.001$, *partial* $\eta^2 = 0.567$. Confidence ratings were higher for sounds presented in the cue condition. None of the other main effects approached significance: vividness, $F(1, 54) = 0.617, p = 0.436$, *partial* $\eta^2 = 0.011$, familiarity, $F(1, 54) = 1.825, p = 0.182$, *partial* $\eta 2=0.033$, and hallucination proneness, $F(1, 54) = 0.454, p = 0.503$, *partial* $\eta^2 = 0.008$. There was one reliable interaction between vividness, familiarity and condition, $F (1, 54) = 9.508, p = 0.003$, *partial* $\eta^2 = 0.150$. In the cue condition, there was no significant difference between low and high vividness stimuli for familiar sounds, $t(55) = 0.503, p=0.617$. Higher confidence ratings were found for high vividness compared to low vividness sounds when the sounds were unfamiliar however, $t(55) = 2.473, p=0.017$. In the no cue condition, participants were significantly more confident in hearing high vividness sounds both when the sounds were familiar, $t(55) = 3.565, p<0.001$, and when the sounds were unfamiliar, $t(55) = 2.987, p=0.004$.

For false positives, the average confidence rating was calculated across vividness, familiarity and condition and a one way ANOVA was computed with hallucination proneness as the between participants factor. For some conditions, participants did not make any false positives; therefore there was a large amount of missing data. The analysis revealed no significant effect of hallucination proneness.

Analysis two: Correlation analysis

Correlations were performed to determine reality monitoring abilities of high and low hallucination prone participants. These analyses also determined how sound detection performance was associated with vividness and familiarity. Here the criteria of $p<0.01$ was used to accept a correlation as significant.

Correlations between sound detection performance and questionnaire measures

The average criterion and d' score was calculated for each participant and these were correlated with the questionnaire scores (LSHS-R, O-LIFE subscales and imagery questionnaires). These correlations were not significant however ($p>0.01$).

Correlations between hits and confidence ratings

To assess how well the two hallucination proneness groups could judge the accuracy of their performance, hit rate and confidence ratings were correlated in each condition. A low correlation would suggest that participants are impaired at judging how accurate they were in the sound detection task (Mintz et al., 1972). Table 23 shows these correlations.

Table 23: Correlation between hits and confidence in hits (r-values) [8]

	Cue	No Cue	Cue (FP)
High hallucination prone	0.746**	0.777**	-0.096
Low hallucination prone	0.795**	0.662**	0.161

**** = p<0.001**

[8] Some sounds received either no hits or no false positives therefore these correlations may differ in the number of items involved in each comparison.

The correlations between hits and confidence ratings were significant at the level of $p<0.001$ for both low and high hallucination prone participants. This showed an association between higher hit rate and increased confidence in responses. Comparison of the strength of these correlations revealed no difference between high and low hallucination prone participants for the cue condition, $z= -0.957$, $p=0.169$. The correlation between hits and confidence ratings was significantly stronger for the high hallucination prone group compared to the low hallucination prone group, $z= 1.908$, $p=0.028$, in the no cue condition however.

Correlations with vividness and familiarity

The following correlation analyses obtained a direct indication of the association between vividness, familiarity and the signal detection measures. The signal detection measures were calculated over item (i.e. for each target sound) for the high and low hallucination proneness groups separately. Each measure was then correlated with vividness and familiarity ratings for each sound obtained from 12 naïve participants from Experiment 3a in Chapter Two.

The measures of response bias and sensitivity could not be calculated for the no cue condition as false alarm responses in this condition were made to a sound category (i.e. high or low vividness) rather than to a specific sound (e.g. cat). Therefore the hit rate was also analysed here in order to reflect these association in the cue conditions. Table 24 shows these correlations for the high and low hallucination prone groups.

Table 24: Correlations between sound detection task measures and vividness, familiarity (r-values)

	Measure	Condition	Vividness	Familiarity
	Hits	Cue	0.205*	0.044
	(N = 128)	No Cue	0.106	-0.017
	FP (N = 128)	Cue	0.004	0.061
	Criterion (N = 128)	Cue	-0.186*	-0.054
	d′ (N = 128)	Cue	0.087	-0.031
	Confidence (hits)	Cue	0.218*	0.059
	(N = 128)	No Cue	0.190*	0.084
High H.P.	**Confidence (FP)** (N = 119)	Cue	-0.086	-0.075
	Hits	Cue	0.123	-0.019
	(N = 128)	No Cue	0.034	-0.116
	FP (N = 128)	Cue	-0.034	-0.015
	Criterion (N = 128)	Cue	-0.024	0.072
	d′ (N = 128)	Cue	0.045	-0.046
	Confidence (hits)	Cue	0.026	-0.083
	(N = 128)	No Cue	-0.161	-0.173
Low H.P.	**Confidence (FP)** (N = 116)	Cue	0.113	0.135

H.P. = hallucination prone
*** = p<0.05**

Vividness ratings: No correlations met the criteria for p<0.01. A number of comparisons met a more liberal criteria of p<0.05, however. For high hallucination prone participants, the hit rate for the cue condition increased for sounds with higher vividness ratings. In addition this analysis revealed that a greater bias was associated with higher vividness ratings. This revealed an association between increased confidence in hearing sounds in both the cue and no cue conditions. For low hallucination prone participants, there was no correlation between vividness rating and any of the sound detection measures.

Familiarity rating: this factor did not correlate with any of the measures, for either the high or low hallucination prone group.

Discussion

The present experiment was based on the hypothesis that imagery vividness, familiarity and hallucination proneness would influence detection of a cued or uncued target sound in white noise. Specifically it was hypothesised that more false positive errors (and therefore a greater bias) would result from high vividness sounds, familiar sounds, and sounds that were cued. In addition, coming from a reality monitoring theory of hallucination development, this study hypothesised that high hallucination prone participants would make more false positive errors, particularly for sounds that were associated with greater imagery vividness. There are three main aspects of this study: i) the effects of imagery vividness and familiarity on sound detection; ii) the effects of the cue on sound detection; and iii) the effect of hallucination proneness on sound detection.

Effects of imagery vividness and familiarity

Participants detected high and low vividness sounds in noise, which varied in familiarity. Imagery vividness and sound familiarity interacted to affect the bias to responding positively to the presence of a sound in noise. When the sounds were familiar, there was no difference in bias between high and low vividness sounds, but when the sounds were unfamiliar, the bias was stronger towards high vividness sounds. Experiment 3, Chapter 2 revealed a similar effect, suggesting a difference in decisional criteria between high and low vividness sounds. There is a lowered response criteria for highly vivid sounds, but this is evened out as the items become more familiar (all high familiarity items having a reduced criterion). This current finding suggests that long-term memory for sounds modulates the imagery vividness effect.

Vividness also interacted with familiarity and condition to influence confidence ratings. In the cue condition, there was no difference between low and high vividness sounds when familiar, but when unfamiliar, participants had greater confidence that they heard high vividness sounds. This mirrors the effect found in the bias data, suggesting that confidence ratings reflect the response bias to respond. In the no cue condition however, participants were generally more confident that they heard high vividness sounds, whether familiar or unfamiliar. This effect is difficult to understand because participants did not know what sound category was presented in this condition. Therefore this suggests that there may have been some differences in the acoustic characteristics of the signals, but not enough to have differential effected sensitivity between high and low vividness sounds (see Appendix H for investigation of the effect of acoustic characteristics of sounds on detection).

Familiarity interacted with condition to affect both the criterion and the d′ measures. In the cue condition, there was no difference in bias or sensitivity between familiar and unfamiliar sounds, but the no cue condition resulted in a greater bias and less sensitivity to unfamiliar sounds. This suggests that the presence of the cue aids the correct detection of unfamiliar sounds, and that, without the cue, the bias is greater for responding positively to unfamiliar sounds. This is probably because the image for such sounds is less defined and therefore its presence is easier to accept. For familiar sounds, detection of the target is much better in the no cue condition, probably because the cue leads to more confusion between the actual target and the image, as these types of sound are high in vividness (see the discussion above). This suggests that the cue has a differential effect on familiar and unfamiliar sounds: a facilitating effect on unfamiliar sound detection and a disruptive effect on familiar sounds.

Effects of the cue

The study also revealed a greater bias towards responding positively in the cue condition compared to the no-cue condition. This suggests that the cue led to participants lowering their criterion to the actual sound present in the noise. Analysis of confidence ratings revealed a similar affect: greater confidence in hearing the sound in the cue condition compared to the no cue condition. The lack of a cue effect on sensitivity contrasts with both Segal and Fusella (1970) (who found that imagining the target impeded sound detection), and Farah and Smith (1983) (who found that imagining the stimulus facilitated detection). The current experiment did not involve explicit instructions to imagine the target, rather just the instruction to ‘listen very carefully’ for the target sound. This differs from the previous studies, which contrasted provision of an instruction to imagine the target and to when instructed not to imagine the target. An issue with the previous studies is that participants always knew what the target sound was and therefore they may have inadvertently imagined the target sound in both the imagery and no imagery conditions (Li, Chen, Yang, Chen, & Tsay, 2002). Evidence that this may occur comes from Merckelbach and van de Ven (2001) who reported that 32% of participants “heard” a song in noise, despite no actual presence of the song. This suggests that people automatically use imagery when exposed to noise conditions, without instruction to do so. Therefore, to resolve the issue of unintended image generation during the no imagery condition, the current study compared performance between when participants knew what the target was (i.e. cue condition) to when they did not know what the target was (i.e. no cue condition).

Effects of proneness to hallucination

Correlation analyses assessed the association between the sound detection measures, vividness, familiarity and acoustic measures for the two hallucination proneness groups. For the high but not the low hallucination prone group, near significant correlations were found between vividness ratings and the hit rate in both the cue and no cue conditions. This suggests that high hallucination proneness may lead to greater sensitivity to high vividness images. The association between vividness, and criterion or d′ for the cue condition was not significant however. This is unusual as the study hypothesised that vividness would have the most effect in this condition, as participants knew what sound they were listening for. This finding may reflect the difficulty in separating images and sounds for high vividness stimuli.

There are several possibilities for why the current study did not find a greater overall effect of hallucination proneness. Firstly, the study used the median score on the LSHS-R to assign participants to the high or low hallucination proneness groups. The issue with this is that there is less separation between the groups compared to, for instance, analysing the extreme scorers (i.e. the top and bottom 25th percentile scorers on the LSHS-R) only, therefore analysis of hallucination proneness in this way may detect fewer differences between the groups. Analysis of the extreme ends of the hallucination proneness scale may have revealed more significant results.

The content of the target stimuli may also have contributed to the lack of a hallucination proneness effect. In patients who hallucinate, the content of the hallucinations are often highly emotionally charged and predominantly verbal (Nayani & David, 1996). The current experiment used verbal material, but as part of the ‘high vividness’ category and was neutral in content. It is possible that cognitive characteristics of the target sounds were not stimulating or emotive enough in the current experiment to reveal differences between high and low hallucination prone participants. Further experiments could manipulate the emotive aspects of the stimuli, to investigate whether such factors influence the hallucination-like experience in the normal population. In addition, research has suggested that hallucinations in schizophrenic patients coincide with affective symptoms, so that hallucination are likely to be more severe when the patient is anxious or depressed (Smith et al., 2006). Future studies should thus look at the effect of current mood or emotional state on sound detection. In addition, future studies could vary the emotional content of the target stimuli too, as previous studies show that hallucinating patients have less vivid imagery for neutral stimuli (Starker et al., 1982) and are more likely to externalise self-generated emotional words than low hallucination prone participants (Laroi, Van der Linden, & Marczewski, 2004).

A further reason for the lack of a hallucination proneness effect is the measure used. The LSHS-R requires participants to read statements about vivid mental events and hallucination-like experience and rate how much each statement applies to them. The problem with this is that it does not take into account how frequently such experiences occur, so that two people may score the same on the questionnaire even though one experiences such events frequently while the other only once. In addition the LSHS-R does not take into account how distracting such experiences are, as it could be hypothesised that such experiences would only cause problems in sound detection tasks if they are highly distracting. The Cardiff Anomalous Perceptions Scale (CAPS) looks more in depth at the types of situations that contribute to hallucinations and has three separate rating scales: distress rating, distraction rating and frequency rating (Bell et al., 2006). Conceivably, this measure could more finely distinguish between high and low hallucination prone participants.

In conclusion this study revealed that little affect of hallucination proneness on sound detection performance. The other factors of sound vividness, familiar and detection cues primarily affected bias towards detecting meaningful sounds in white noise, and confidence in those responses however. Familiarity also affected sensitivity to the target, as participants were more sensitive to familiar sounds. These results therefore suggest that vividness, familiarity and cues have top-down influences on sound detection performance.

Chapter Six. Emotional Words, Auditory Imagery Vividness and Sound Detection

Abstract

Psychiatric auditory hallucinations are often related to emotional state and can be emotional charged. Experiment 7a determined whether auditorily-presented emotional words elicit more vivid auditory images and generate better memory recall. Experiment 7b employed the same sound detection paradigm as in Chapters 2 and 5 and compared detection of emotional words in noise differs to detection of neutral words. Both studies compared high and low hallucination prone participants on these measures. Experiment 7a found that participants rated emotional words as more vivid and had a higher recall rate than neutral words. In Experiment 7b emotional words resulted in a lower response bias and increased sensitivity compared to neutral words. Response bias increased and sensitivity decreased when participants received a cue to the target sound, particularly for emotional compared to neural words. In addition associations between performance on the sound detection task and self-reports of anomalous experiences were reliable. The current study finds an association between auditory imagery and emotional material, and shows that hallucination proneness results in greater sensitivity, confidence and word recall with verbal material, though it appears not to interact with emotion. Reasons for these findings are discussed.

Experiment 7. Emotional Words, Auditory Imagery Vividness and Sound Detection

Introduction

Auditory Verbal Hallucination (AVH) is a key symptom for diagnoses of schizophrenia (Beck et al., 2003) however the mechanisms by which hallucinations develop are not fully understood. Many authors posit that there is disruption of auditory verbal imagery, whereby internally generated speech is misattributed to be from an external source (McGuire et al., 1996a). Also it has been suggested that vivid imagery and hallucinations exist on either end of the same continuum (Slade & Bentall,1988).

As mentioned in previous chapters however, the link between vividness of imagery and experience of hallucinations is a far from clear (Aleman et al., 2001; Bocker et al., 2000; Sack, van de Ven, Etschenberg, Schatz, & Linden, 2005). Rather than linking auditory hallucinations directly to changes in the strength of auditory images, other theorists have argued that the problem lies in the attributions of those auditory images. Brebion et al. (2002) found that hallucinating patients are more likely to attribute self-generated items to an external source in a recollection task. Patients gave related words to a list of probe-words (e.g. cat – dog). They then saw the items they first received (i.e. old items), their own related words (i.e. self generated new items) and other related words (i.e. experimenter-generated new items) and asked if the stimuli were old or new items. If new, participants also indicated whether they were self- or experimenter-generated items. High hallucination prone participants were more likely to attribute their own words as being generated by the experimenter, indicating a general source monitoring bias to external rather than internal sources.

Irrespective of the cognitive processes linked to auditory hallucinations, one attribute of these hallucinations is that they are often emotionally charged (Nayani et al., 1996). Disturbance of emotional experience and emotion recognition is a key feature of schizophrenic disorders (Kraeplin, 1907). Hallucinating patients make more errors in judging emotion from the prosody of a voice and from a facial expression, compared with control participants (Edwards et al., 2002). Alteration in the emotional connotation of an image may also be a contributory factor to AVHs, along with alterations in image quality and/or attributions (Kerns, 2005). Consistent with this are studies demonstrating effects of emotion on source monitoring. Laroi et al. (2004), for example, argued that source monitoring errors increased for emotional material because emotional content disturbs normal encoding. They employed a similar task to that of Brebion et al. (2002) except that positive, negative and neutral stimuli were used as word-probes. High hallucination prone participants made more source discrimination errors, attributing self-generated items to the experimenter, particularly for emotionally charged material, and not for neutral material. This suggests that the emotional content of the stimuli exacerbated the source monitoring deficits of the high hallucination prone group.

Relatively little previous research focused on the effects of emotion on auditory perception tasks with high and low hallucination prone individuals. Tests of auditory perception may more closely resemble conditions under which hallucinations might arise however (e.g. with perceptual input occurring in noise) such as the tasks used by

Bentall and Slade (1985) and Barkus et al. (2007) (see Chapter 1 for a review of these studies). In the majority of cases however, experimenters have used neutral sounds such as tones, whereas hallucinations are mainly verbal, and emotive in content. Studies using neutral sounds may be less sensitive to effects of hallucination-proneness on auditory imagery and sound detection (Li et al., 2002; Morrison et al., 1997).

The present study assessed the relation between proneness to hallucination and both auditory imagery and auditory perception using words varying in their emotional content. Experiment 7a compared high hallucination-prone participants to low hallucination-prone participants on a task involving auditory imagery ratings for positive, negative and neutral words. The study hypothesised an association between the emotionality of the words and increased auditory imagery vividness, and that this association would differ between the two hallucination proneness groups. Baddeley and Andrade (2000) found that long term memory factors such as meaningfulness and familiarity can affect imagery vividness ratings. The effects of emotional valence on auditory imagery have not been studied before though, and no studies have examined whether hallucination proneness moderates any of the effects of emotional valence.

Experiment 7b compared high and low hallucination prone participant in their bias and sensitivity to positive, negative and neutral words presented in white noise. It was hypothesised that there would be more source monitoring confusions in sound detection by high hallucination prone participants. If so, then there would be an increased bias to detect emotional words and (possibly) decreased sensitivity. Experiment 7b also compared performance in two conditions: (i) a cue condition where participants knew the target words and imagined the words during detection; (ii) a no cue condition where participants had no knowledge of the target words beforehand. It was hypothesised that source monitoring errors would be greatest in the cue condition for emotional words and for high hallucination prone participants.

As well as assessing auditory imagery and testing auditory word detection, we also examined memory for the different word types in Experiment 7a by giving participants a surprise recall task at the end of their rating study. Experiments with normal participants suggest that memory for emotional items is better than that for neutral items (Danion, Kazes, Huron, & Karchouni, 2003), and that people tend to remember emotional events more vividly (Kensinger & Corkin, 2003). Following this, we may expect better recall for emotional words than neutral words. The study assessed whether there is an exaggeration of this affect in individuals prone to hallucinations.

Experiment 7a. Auditory imagery vividness of emotional words.

Method

Participants

24 participants from the University of Birmingham completed this experiment (2 male, 22 female, mean age: 19.79 years). Each participant completed the Revised Launay-Slade Hallucination Scale (LSHS-R; Bentall et al., 1985). Participants were classified as high or low hallucination prone, according to whether they scored high or low on this scale. All participants received credit in return for their participation. Given that the participants were all part of a single undergraduate cohort, medication,

current cognitive state, age or IQ level is unlikely to confound performance (Aleman et al., 2001).

Design

Participants rated their imagery vividness for emotional and neutral words. Once they rated the vividness for each word, participants completed a surprise memory test for the stimuli words.

Stimuli

The stimuli were selected from the ANEW (Affective Norms for English Words; Bradley & Lang, 1999). Thirty positive, 30 negative and 30 neutral words were chosen from this list, and a recording of each word was taken, spoken by an English female speaker, in a neutral voice. These words were matched for frequency across emotional categories. Each recording lasted approximately 1sec.

Procedure

Appendix F shows the instructions given to the participants about the task. On each trial participants listened to a word, followed by 5sec of white noise. A rating scale then appeared on the screen. Participants rated how clearly they could imagine the spoken word, on a scale of one to five (one being “no image” and five being “image is clear and vivid”). Reaction times to make these ratings were also recorded.

Each word was presented once and when the experiment finished, the participants wrote down as many words as they could remember from the experiment.

Results

Questionnaire

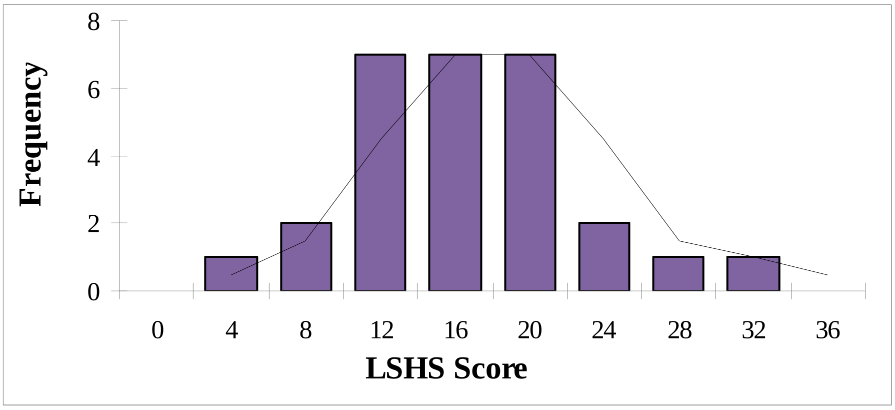


Figure 25: Distribution of LSHS-R scores

Figure 25 shows the distribution of the LSHS-R scores. The average score was 15.29 (SD = 6.39). Participants were assigned to a hallucination proneness group, according to the median value on this questionnaire. The median score was 14, therefore participants scoring above this score were grouped as the high hallucination prone group and those scoring below this score were the low hallucination prone group

Imagery vividness

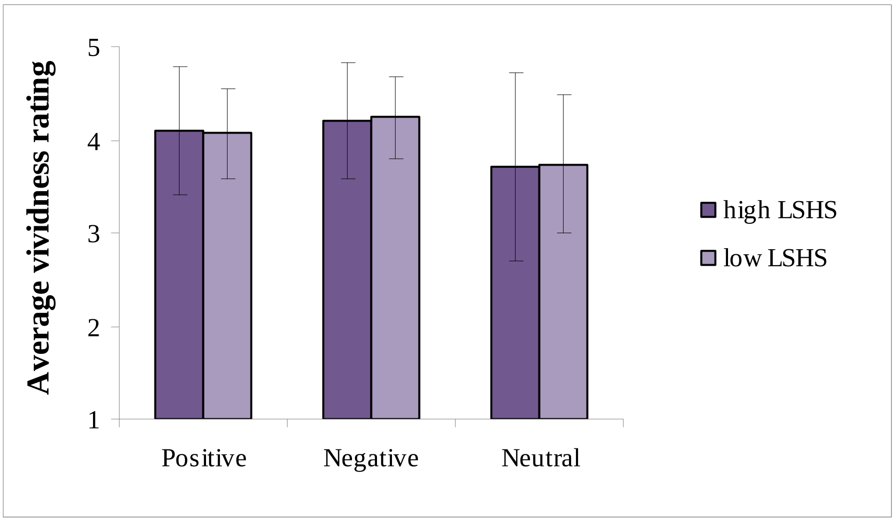


Figure 26: Average vividness rating for high and low hallucination prone participants, for each emotion

The average vividness rating for each emotion was calculated for each participant (see Figure 26). A 3 x 2 split plot ANOVA was used to analyse imagery vividness ratings, with word emotion as the within participants factor and hallucination proneness as the between participants factor. This revealed a significant main effect of emotion $F(2, 44) = 19.528$, $p < 0.001$, partial $\eta 2 = 0.470$. Bonferroni adjusted post hoc analyses revealed that vividness ratings were significantly higher for negative and for positive words compared to neutral words (both $p<0.001$). There was a borderline difference between negative and positive words ($p=0.053$), as negative words received higher ratings. There was no significant main effect of hallucination proneness $F(1, 22) = 0.026$, $p = 0.874$, partial $\eta 2 = 0.001$, and no interaction between emotion and hallucination proneness $F(2, 44) = 0.157$, $p = 0.758$, partial $\eta 2 = 0.007$.

RT to make vividness rating

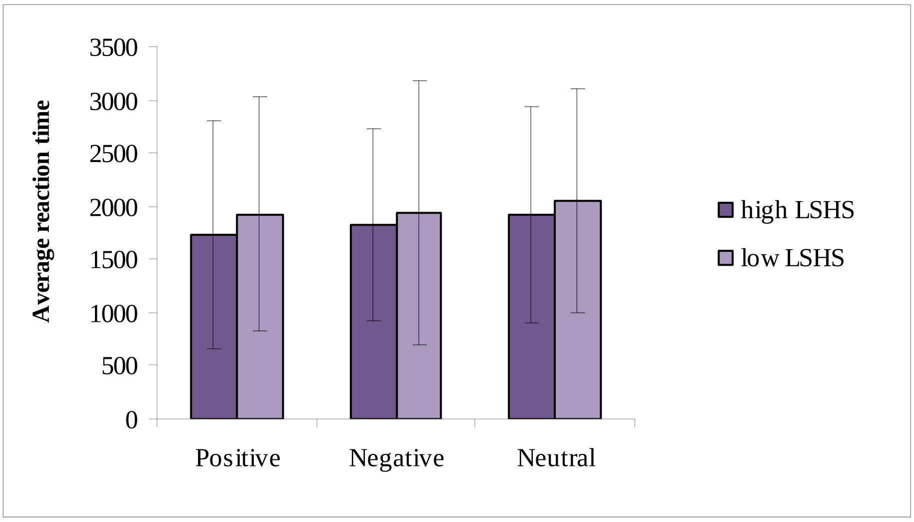


Figure 27: Average reaction time for high and low hallucination prone participants for each emotion

The average reaction time for each emotion condition was calculated for each participant (see Figure 27). The data were analysed using the same design ANOVA as for the rating results. There were no significant main effects or interactions.

Memory

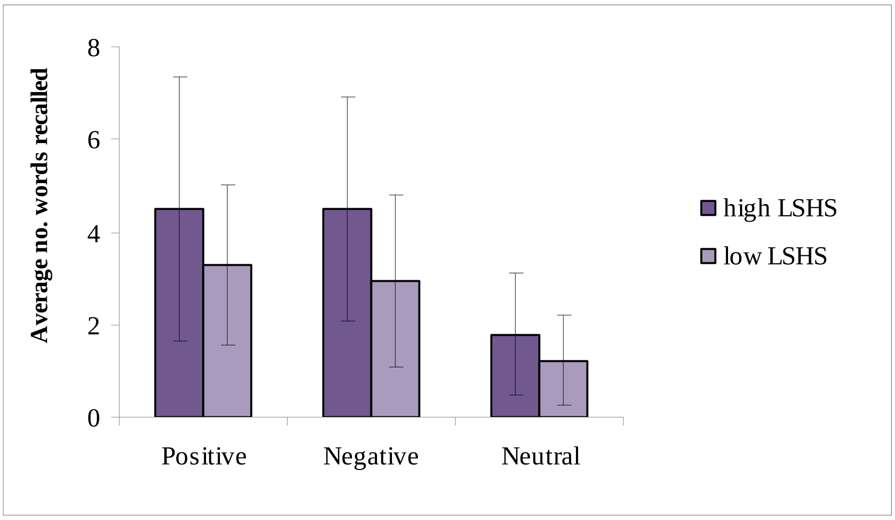


Figure 28: Average number of words recalled for each emotion condition by high and low hallucination prone participants

The average number of words recalled in each category was calculated for each participant (see Figure 28). A two-factor split plot ANOVA revealed a significant main effect of emotion, $F(2, 44) = 15.641, p < 0.001$, *partial* $\eta 2 = 0.416$. Bonferroni adjusted post hoc analyses revealed that significantly more negative words were recalled than neutral words ($p<0.001$) and more positive words were recalled than neutral words ($p=0.002$). There was no significant difference between the number of negative and positive words recalled ($p=0.632$) however. There was also a significant main effect of hallucination proneness, $F(1, 22) =4.624$, $p = 0.043$, *partial* $\eta 2 = 0.174$; high hallucination-prone participants recalled more words than low hallucination-prone participants. The interaction between emotion and hallucination-proneness was not significant, $F(2, 44) = 0.655$, $p = 0.468$, *partial* $\eta 2 = 0.029$[9].

[9] Further analyses assessed if memory performance varied as a function of whether the words appeared early, in the middle or in the final set of stimuli. However, the serial position of the words in the rating study had no reliable impact on recall and this factor did not interact with any of the other variables of interest.

Correlation analyses

The average imagery rating and the proportion of participants who recalled each word was calculated for the low and high hallucination-prone groups. The correlation between vividness and memory was significant for both high hallucination-prone participants, *r (90) = 0.264, p = 0.012,* and low hallucination-prone participants, *r (90) = 0.335, p<0.001*, revealing that participants recalled more of the words that they rated higher in imagery vividness.

The vividness rating for each word and the proportion of participants who recalled each word was also correlated with word frequency, valence rating and the arousal rating (ratings from ANEW; Bradley et al., 1999). Table 25 shows these correlations. These analyses suggest that the relations between imagery vividness and memory were modulated by arousal.

Table 25: Correlations for vividness and memory recall with frequency an arousal rating (r-values)

Correlation		**High LSHS-R**	**Low LSHS-R**	**Total**
Vividness	**Freq.**	0.107	-0.072	-0.100
	Arousal	0.482**	0.565**	0.583**
	Valence	-0.148	-0.090	-0.133
Memory	**Freq.**	0.058	0.035	0.050
	Arousal	0.318*	0.253*	0.311*
	Valence	-0.001	0.063	0.036

N = 30; * - p<0.05; ** - p<0.001

Arousal rating was used as a covariate in a correlation between vividness and memory score. The correlation between these two variables disappeared for both high hallucination prone participants, *r(78) = 0.105, p=0.352* and low hallucination prone participants, *r(78)= 0.211, p= 0.061*[10].

ANCOVA was then used to assess the effect of arousal on vividness and memory Vividness ratings and memory scores were analysed over items using a split plot ANOVA, with arousal rating as a covariate. The within-participants factor in these analyses was hallucination proneness (high and low hallucination prone participants) and the between participants factor was emotional content (positive, negative and neutral).

For the vividness ratings, without the covariate there was a significant main effect of hallucination proneness, $F(1, 87) = 45.459$, $p < 0.001$, partial $\eta 2 = 0.343$; the high hallucination-prone participants rated their auditory imagery as being significantly higher than the low hallucination-prone participants. With arousal rating as a covariate, the main effect of hallucination proneness, $F(1, 77) = 1.787$, $p = 0.185$, partial $\eta 2 = 0.023$, disappeared. In addition without the covariate the main effect of emotion was significant, $F(2, 87) = 32.013$, $p < 0.001$, partial $\eta 2 = 0.424$, and Bonferroni adjusted post hoc analysis revealed significantly higher ratings for negative and positive words compared to neutral words (both $p<0.001$) but no significant difference between negative and positive words ($p=0.151$). Again, with the covariate of arousal rating, this main effect disappeared, $F(2, 77) = 2.103$, $p = 0.129$, partial $\eta 2 = 0.052$. The interaction between hallucination proneness and emotion was non- significant both with and without the covariate.

[10] Arousal ratings were unavailable for 9 neutral words, as these came from another normed list (MRC Psycholinguistic Database; Coltheart, 1981)

For the memory scores, without the covariate there was no significant main effect of hallucination proneness, $F(1, 87) = 2.095$, $p = 0.151$, partial $\eta2 = 0.024$, which remained non-significant with addition of the arousal covariate, $F(1, 77) = 0.310$, $p = 0.580$, partial $\eta2 = 0.004$. The main effect of emotion was significant without the covariate, $F(2, 87) = 3.972$, $p = 0.022$, partial $\eta2 = 0.084$. Bonferroni adjusted post hoc analysis revealed a trend for recall of negative words being higher than that for neutral words ($p=0.091$); also more positive words were recalled than neutral words ($p=0.031$) but there was no significant difference in the number of positive and negative words recalled ($p=1.00$). The inclusion of arousal as a covariate removed this significant effect, $F(2, 77) = 0.318$, $p = 0.729$, partial $\eta2 = 0.008$. The interaction between hallucination proneness and emotion was non- significant both with and without the covariate.

Discussion

This experiment examined how the emotional connotation of a word influences auditory imagery and memory. Emotionally positive and negative words received higher imagery vividness ratings than neutral words, and the two types of emotion words did not differ. There was also a significant correlation between arousal rating for each word and the vividness ratings, but not between vividness ratings and frequency. This therefore suggests that increased arousal contributes to increased ability to imagine spoken words, independently of the overall frequency of the word. The fact that use of arousal rating as a covariate removed the main effect of emotion confirmed this.

Experiment 7a also assessed memory for the words used in this task by using a free recall task. This revealed that participants recalled more of the positive and negative items compared to the neutral items. In addition there was a significant correlation between auditory imagery vividness and recall. This experiment therefore supports Kensinger and Corkin's (2003) finding of increased memory for emotional items, and extends it to include auditory imagery vividness.

Interestingly hallucination proneness had little effect on imagery, but there was an effect on memory. High hallucination prone participants remembered more items from the vividness experiment, but this did not interact with the emotional class of the stimulus. Overall differences in memory performance between high- and low hallucination-prone individuals have not been explored before, but at least one component of the observed difference may reflect the arousal-inducing capacity of the words. The negative and positive emotion words were associated with high arousal, and it may be that high hallucination prone participants have a greater arousal response to such stimuli. Memory might improve overall if participants are in a more aroused state. Indeed correlation between arousal rating and memory score were significant for both hallucination-proneness groups, but stronger so for high hallucination prone participants. Adding the differences in arousal ratings as a covariate to the analyses removed the effect of emotion on memory scores; hallucination proneness was non-significant, both with and without the arousal rating.

There was no effect of hallucination proneness on imagery vividness ratings. One reason for this is that participants were grouped according to whether their score on the LSHS-R fell above or below the median, rather at the extreme ends of the scale.

The data were therefore reanalysed with just the participants who scored at the top and bottom 25th percentile, however hallucination-proneness remained non-significant. This may simply be due to the subsequent groups having a relatively small number of participants, and the study may have found significant effects with a larger sample of extreme scorers

The item analyses provided support for the effect of hallucination-proneness on imagery and memory. There was a reliable increase in imagery vividness ratings for the high hallucination-prone group even with the use of a median split approach. Interestingly this effect disappeared with arousal rating as a covariate, suggesting that arousal contributes somewhat to the difference in vividness ratings between high and low hallucination prone participants. If the high imagery vividness of high hallucination-prone individuals reflect real differences in imagery, then it is possible that a contrasting pattern of performance will emerge with the high- and low hallucination prone individuals on an auditory detection task. For example, there may be a tendency for high hallucination-prone individuals to make false positive errors when cued to hear signals, if they have difficulty discriminating heard from imagined sounds. This could also lead to changes in sensitivity. One possibility is that, if the stimulus signal is added-to by the auditory imagery for the stimulus, then this may enhance sensitivity (i.e. as suggested by Farah et al., 1983). It is also possible however, that difficulty in discriminating between the signal and the imaged stimulus could decrease sensitivity (i.e. as suggested by Segal et al., 1970). Experiment 7b tested these possibilities. In this experiment the participants heard the emotional (positive and negative) and neutral words in noise and decided on the presence of a word in the noise on each trial. The words were also either cued or not-cued. The effects of imagery on performance should be particularly strong when words are cued.

Experiment 7b. Emotional words and auditory word detection

Method

Participants

Nine male and 43 female participants from the University of Birmingham volunteered for this experiment (mean age: 21.27 years) in exchange for psychology course credits. All had self-reported normal hearing.

Design

The current experiment required participants to listen to 3sec clips of white noise, and indicate whether or not they heard one of the spoken words embedded in the noise. If the word was present, it was presented in the middle portion of the white noise. Each speech clip lasted 1sec, spoken by a female native English speaker. Participants detected the words under two conditions. In the cue condition the spoken word was presented prior to the detection period. Each word occurred once at the audible level, once at the threshold level and on two occasions the word was cued, but not presented. Therefore there were a total of 120 trials. In the no cue condition, the participant did not know what word they were listening for. Again 120 trials were presented. Each word was presented once at an audible level, once at the threshold level, and the remaining 60 trials featured just white noise.

The experiment was presented using the experiment design program, E-Prime (Psychology Software Tools, Inc.) and auditory stimuli were presented over Sennheiser 250 headphones.

Stimuli

Ten positive, negative and neutral words were selected from those used in the previous experiment. These sounds were then manipulated, so as to be at an audible level and a threshold level when presented in white noise. The threshold level was established in an earlier experiment, as the point at which participants detect between 70 and 75% of the sounds when presented in white noise.

Questionnaires

All participants completed the LSHS-R (see previous experiment for description), and were grouped according to whether they scored above or below the median score.

In addition, 39 participants also completed the Cardiff Anomalous Perceptions Scale (Bell et al., 2006). Participants gave yes or no responses to 32 questions about unusual sensory and imaginary experiences in a range of modalities. In addition if participants gave a positive response they also rated the experience on three subscales: intrusiveness of the experience, distress caused by the experience and frequency of the experience. Therefore each participant had four scores on this scale, i.e. the total score, and the three subscale scores.

Procedure

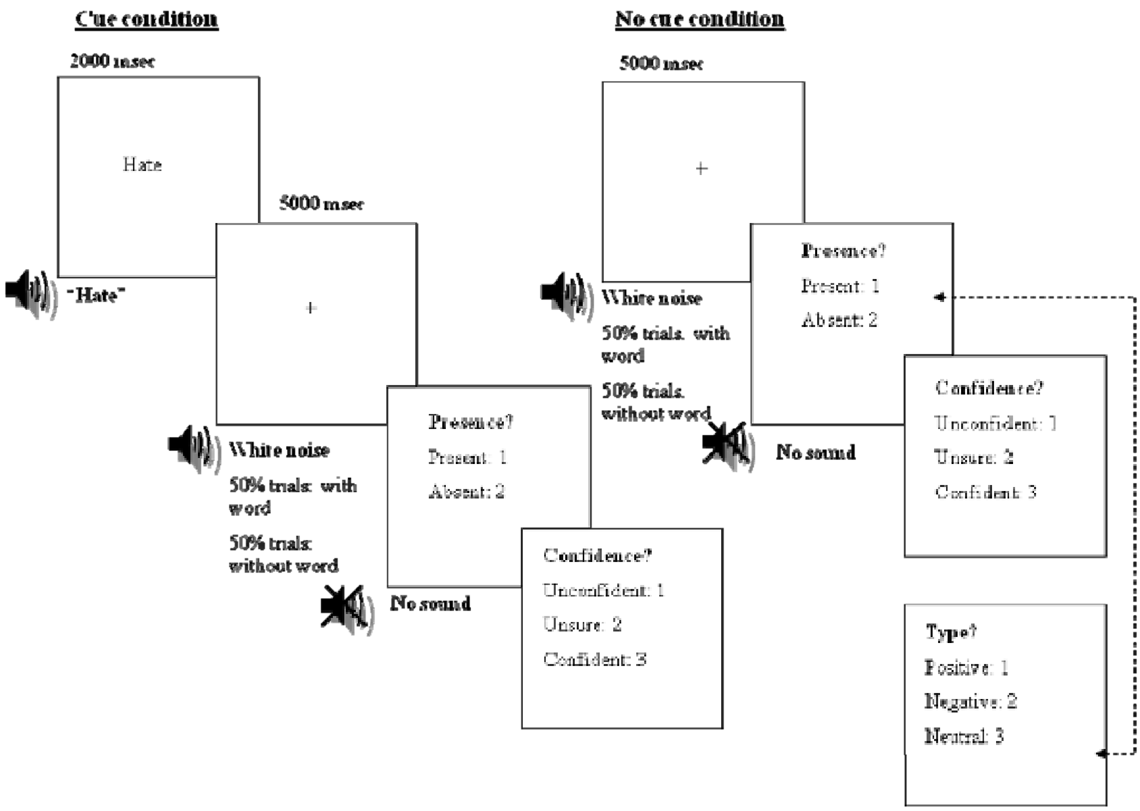


Figure 29: Experiment procedure in each condition

Appendix G shows the written instructions that participants received. The experimenter first informed participants that they had to detect spoken emotional and non-emotional words in white noise. The experimenter then told the participants that the word may or may not be present in the white noise, but that if it was present, it would be either at an audible level or ay a threshold level. Therefore they were asked to listen very carefully for the words. In addition they were told that on half of the trials they would receive a cue as to the word they were listening out for. Participants were told to hold the word in memory and that for the remaining trials they would receive no cue about the word.

Figure 29 shows the procedure of each trial. At the beginning of the trial, participants saw a condition prompt (i.e. “cue” or “no cue”). In the cue condition, participants then heard the word that they were listening for. A fixation cross subsequently appeared on the screen for 1sec, followed by 3sec of white noise. Following this, participants the indicated whether the cued word was present in the white noise or was absent, by pressing either “1” or “2” on the computer keyboard. Subsequently there was a prompt asking participants to indicate how confident they were in their answer, by rating from 1 (being unconfident) to 3 (being confident) on the computer keyboard. In the no cue condition, participants did not receive a sound cue, rather they just heard white noise, and were again prompted to indicate whether the sound was present or absent, and how confident they were in this answer. In addition, if the participant indicated that they did hear a word, they were asked to indicate whether it was a positive, negative or neutral word that was heard, by pressing a number on the keyboard.

The participants received four practice trials, prior to commencement of the full experiment, which lasted approximately 40 minutes. At the end of the experiment, participants completed the Launay-Slade Hallucination Scale, and a subset also completed the CAPS

Results

Five participants were excluded from the experiment, as their overall average proportion of detected words fell below 50%. Therefore data from 47 participants was analysed.

Questionnaires

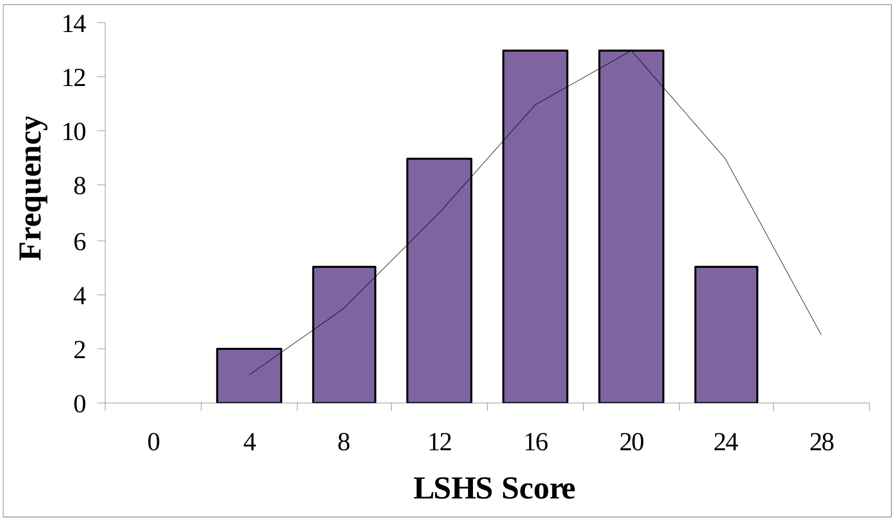


Figure 30: Distribution of LSHS scores

Figure 30 shows the distribution of the LSHS-R scores. The mean score on this questionnaire was 14.19 (SD=5.25). Participants were then grouped as high or low hallucination prone according to whether their LSHS-R score fell above or below the median value on the LSHS-R. Twenty-four participants scored on or above the median score of 16 (six participants scored the median value) and were therefore included in the high hallucination prone group. Twenty-three participants scored below the median and were included in the low hallucination prone group.

Table 26: Average and standard deviations of the CAPS subscale scores

	Total Score	**Distress Score**	**Intrusiveness Score**	**Frequency Score**
Mean	6.56	12.69	14.87	11.18
SD	3.94	9.82	10.96	7.04

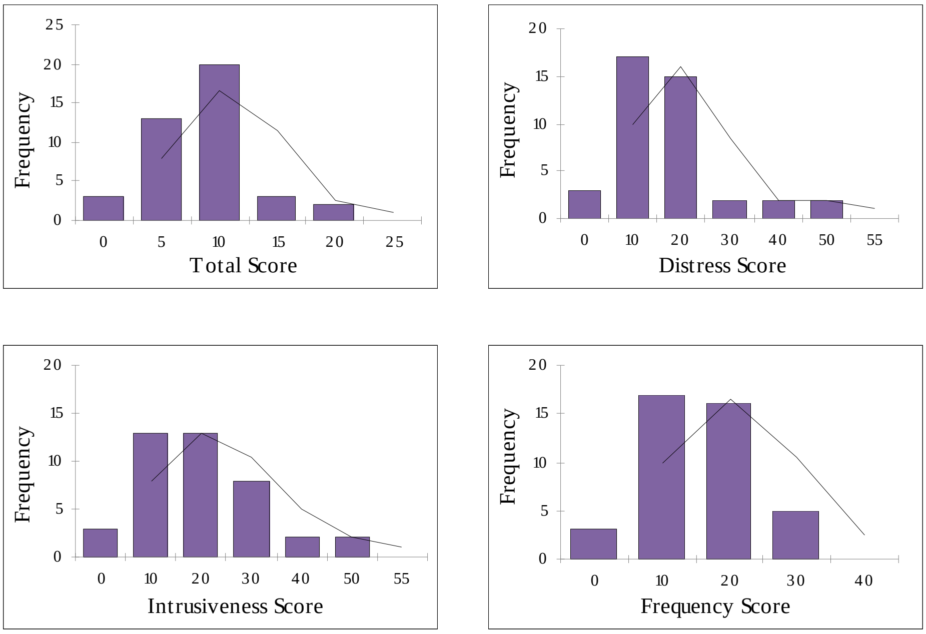


Figure 31: Distribution of CAPS subscale scores

Table 26 shows the average scores and standard deviations on each CAPS subscale, and Figure 31 shows the distribution of these scores.

Sound Detection Task

The data were analysed at three levels. The first analysis focused on the sound detection measures of response bias (c) sensitivity (d') and confidence rating. Repeated measures ANOVA’s were performed on these measures to determine the affect of emotion, hallucination proneness and cue condition on detection of the words in noise. The second analysis focussed on correlations between the different measures. Hits and confidence ratings were correlated to determine reality monitoring abilities of the two hallucination proneness groups. In addition the measures from the sound detection task were correlated with other measures associated with the target words, such as vividness, frequency and arousal. This analyses determined how such factors influence sound detection, and whether there is a differential relationship between them, based on hallucination proneness. The third analysis focussed on the CAPS questionnaire scores, to determine whether other anomalous experiences (apart from auditory hallucinations) are related to performance in sound detection tasks.

Figure 32 shows a plot of hits against false alarms in each condition. The plot suggests a close relationship between most conditions in hit and false alarm rates, with high hallucination prone participants in the cue condition appearing to have the highest hit and false alarm rate, particularly for negative stimuli.

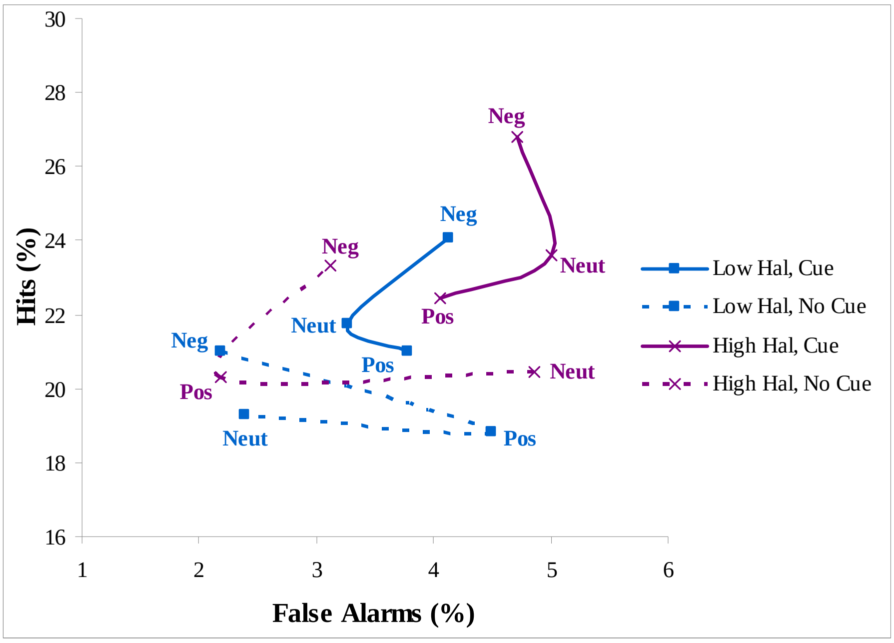


Figure 32: Plot of hits against false alarms in each condition

Analysis one: Effects of emotion, hallucination proneness and cue condition on sound detection

As in Chapters 2 and 5, the hits and false positive rates for each participant in each condition were converted into response bias (c) and sensitivity (d′) measures, using the method noted by Stanislaw (1999) described in Chapter 2. Analysis of these measures and the confidence rating scores consisted of a series of split plot ANOVAs.

Criterion

Table 27: Average criterion for high and low hallucination prone participants (standard deviations in brackets)

		Positive	Negative	Neutral
Cue	High H.P	1.41 (0.54)	1.26 (0.41)	1.44 (0.47)
	Low H.P	1.58 (0.54)	1.41 (0.44)	1.49 (0.40)
No Cue	High H.P	1.75 (0.56)	1.62 (0.58)	1.47 (0.52)
	Low H.P	1.76 (0.52)	1.76 (0.53)	1.60 (0.52)

The criterion assesses each participant's bias towards responding with a positive answer. The lower the criterion, the greater the participants bias towards responding positively in an ambiguous signal plus noise environment and the greater the number of false positive responses made. Table 27 shows the average criterion for each condition, for high and low hallucination prone participants.

A three-way split plot ANOVA (within-participant factors: emotion and condition, between participant factor: hallucination proneness) revealed a main effect of emotion, $F(2, 90) = 5.516, p = 0.006, \text{partial } \eta 2 = 0.109$. Bonferroni adjusted post hoc analyses revealed that negative words had a significantly lower criterion than positive words (p=0.019), and neutral words had a significantly lower criterion than positive words (p=0.020). The significant main effect of condition revealed a greater bias for words in the cue condition, $F(1, 45) = 20.085, p < 0.001$, $\text{partial } \eta 2 = 0.309$. There was no significant main effect of hallucination proneness, $F(1, 45) = 0.849, p = 0.362, \text{partial } \eta 2 = 0.019$.

Finally there was an interaction between emotion and condition, *F(2, 90) = 4.194, p = 0.021, partial η2 = 0.085*. There was a lower criterion in the cue compared to the no cue conditions for negative words, *t(46) = 4.933, p<0.001*, and for positive words, *t(46) = 3.024, p=0.004*, but no such difference for neutral words, *t(46) = 1.037, p=0.305*.

D′

Table 28: Average d' for high and low hallucination prone participants (standard deviations in brackets)

		Positive	**Negative**	**Neutral**
Cue	High H.P	1.38 (0.97)	1.28 (0.72)	1.35 (0.89)
	Low H.P	1.53 (0.91)	1.41 (0.80)	1.40 (0.78)
No Cue	High H.P	1.83 (0.99)	1.75 (0.96)	1.31 (0.96)
	Low H.P	1.75 (1.01)	1.88 (0.93)	1.40 (1.05)

D′ is a measure of sensitivity to a signal. A high d′ prime indicates a higher sensitivity to the target sound. Table 28 shows the average d' for each condition for the high and low hallucination prone group.

A three way split plot ANOVA conducted on the d′ data revealed a significant main effect of emotion, *F(2, 90) = 5.751, p = 0.005, partial η2 = 0.113*. Bonferroni adjusted post hoc analysis revealed greater sensitivity for negative sounds compared to neutral sounds *(p=0.033)* and greater sensitivity for positive words compared to neutral words *(p=0.017)* but no difference between negative and positive words (*p>0.05*). There was also a significant main effect of condition, *F(1, 45) = 6.572, p = 0.014, partial η2 = 0.127*; sensitivity was greater in the no cue condition. The effect of hallucination proneness was not significant, *F(1, 45) = 0.148, p = 0.702, partial η2= 0.00*3, but there was an interaction between emotion and condition, *F(2, 90) =*

3.437, p = 0.045, partial η2 = 0.071. This revealed greater sensitivity to negative words in the no cue condition compared to the cue condition *F(1, 46) = 12.440, p = 0.001, partial η2 = 0.213*. This effect approached significance for positive words *F(1, 46) = 3.845, p = 0.056, partial η2 = 0.077*, but there was no significant difference between the conditions for neutral sounds *F(1, 46) = 0.019, p = 0.891, partial η2 < 0.001*.

Confidence Ratings

Table 29: Average confidence rating for high and low hallucination prone participants in each condition

		Positive	Negative	Neutral
Cue	High H.P	2.66 (0.26)	2.80 (0.20)	2.74 (0.25)
	Low H.P	2.72 (0.27)	2.78 (0.20)	2.76 (0.20)
No Cue	High H.P	2.57 (0.29)	2.66 (0.36)	2.58 (0.39)
	Low H.P	2.49 (0.41)	2.50 (0.38)	2.62 (0.37)

Table 29 shows the average confidence in correct responses, for each condition. A three-way split plot ANOVA analysed the confidence ratings for hits (within-participants factors: emotion and condition; between participants factor: hallucination proneness).This revealed a trend towards a main effect of emotion, *F(2, 90) = 3.035, p = 0.059, partial η2 = 0.063*. Confidence was highest when the word was negative, followed by neutral words and positive words, though Bonferroni adjusted post hoc comparisons were not significant (*p>0.05*). There was also a significant main effect of condition, *F(1, 45) = 19.986, p<0.001 , partial η2 = 0.308*, revealing that confidence was greater in the cue condition, compared to the no cue condition. No other main effects or interactions were significant.

A three-way split plot ANOVA also assessed the confidence ratings in false positive (within-participants factors: emotion and condition; between participants factor: hallucination proneness). Because some participants did not make any false positive in one or more of the conditions, this analysis was based on 14 high hallucination participants and 17 hallucination prone participants. There were no main effects or interactions.

Confidence ratings for the false positives were analysed in a similar way, however this revealed no significant main effects or interactions.

Analysis two: Correlational analysis

These analyses determined how performance on the sound detection task was associated with hallucination proneness, imagery vividness and characteristics of the words (i.e. arousal, valence and frequency ratings). In addition a further analysis determined reality monitoring abilities of high and low hallucination prone participants. Due to the number of comparisons a stringent criteria of $p<0.01$ was used to accept significance.

Correlation between sound detection measures and hallucination proneness

Correlations between the LSHS-R score and the average criterion, d′ and confidence ratings, aimed to assess the existence of a relationship between these factors. Neither the association between LSHS-R and the average criterion adopted by participants, nor their average d′ score were significant (criterion: $r(47)=0.114$, $p=0.334$; d′: $r(47)=0.089$, $p=0.553$). The correlation between LSHS-R and the average confidence rating to hit and false alarm responses was also not significant, (hits: $r(47)=-0.155$, $p=0.297$; false positives: $r(47)=-0.145$, $p=0.435$).

Correlation between hits and confidence rating

Correlation between the average hit rate and the average confidence rating for each item for the high and low hallucination prone groups established reality monitoring abilities. For the cue condition, this analysis revealed significant correlations for both high hallucination prone participants, $r(30)=0.724$, $p<0.001$, and low hallucination prone participants, $r(30)=0.891$, $p<0.001$. The correlation was significantly higher for the low hallucination prone group, $z=1.877$, $p = 0.030$.

These correlations were significant in the no cue condition for low hallucination prone participants, $r(30)=0.683$, $p=0.001$, but did not meet the criteria of $p<0.01$ for high hallucination prone participants, $r(30)=0.391$, $p=0.032$. Comparison between these correlations revealed a borderline higher correlation for low hallucination prone participants difference, $z= 1.55$, $p = 0.060$.

Correlations with vividness, arousal, valence, word frequency and rating

The vividness rating for each sound in the current experiment was taken from the previous experiment, along with the frequency, arousal and valence rating of each word (rating from ANEW; Bradley et al., 1999). The measures were correlated with the hit and false alarm rate, confidence rating, criterion and d′ for each word.

There was a significant correlation between vividness and arousal rating, $r(30) = 0.679$, $p<0.001$, showing that vividness rating was higher for words that were higher in arousal rating. There was also a near significant correlation between valence rating and frequency $r(30) = -0.409, p=0.025$; words lower in valence (i.e. more negative in emotion) were more frequent. No other correlations were significant ($p>0.01$).

The correlations between vividness rating and the criterion did not reach the criteria of $p<0.01$, for low hallucination prone participants, $r(30) = -0.023, p=0.905$. The association approached significance for high hallucination prone participants, $r(30)=-0.335, p=0.070$. The correlation between valence rating and confidence ratings on false positive trials was significant for high hallucination-prone participants, $r(30) = 0.467, p=0.009$. These participants were more confident that they heard words with a higher valence (i.e. more positive in emotion). This correlation was not significant for the low hallucination prone participants, $r(30) = -0.041, p = 0.829$

Analysis three: Association between anomalous experiences and sound detection performance

Correlations with CAPS scores

In order to assess whether these factors influenced sound detection, each of the four CAPS scores (total, distress, distraction and frequency scores) was correlated with the criterion and d′ scores in each condition. In addition, within the questionnaire there were nine categories related to different anomalous experiences. The number of questions applying to these categories varied, and some questions applied to more than one category. These categories were:

1. Changes in levels of sensory intensity (5 items)
2. Having a non-shared sensory experiences (4 items)
3. Inherently unusual or distorted sensory experiences (5 items)
4. Sensory experience from an unexplained Source (6 items)
5. Distortion of form of own body and of external world (4 items)
6. Verbal hallucinations (3 items)
7. Sensory flooding (2 items)
8. Thought echo and hearing thoughts out loud (2 items)
9. Temporal lobe disorder (4 items)

Subscale correlations

Table 30: Correlation between total scores on CAPS subscales and signal detection measures

	Criterion			d′		
	Positive	**Negative**	**Neutral**	**Positive**	**Negative**	**Neutral**
Total	0.112	0.205	0.205	0.085	0.210	0.205
Distress	0.106	0.173	0.157	0.080	0.144	0.111
Distraction	0.095	0.192	0.203	0.053	0.171	0.171
Frequency	0.194	0.296	0.238	0.190	0.268	0.243

N = 39

Correlations between the CAPS subscales and the signal detection measures were not significant.

Category correlations

Total score: Though no correlations reached our stringent criteria of p<0.01, two weaker correlations were found. The 'Changes in Sensory Intensity' subscale correlated with the criterion and d′ for neutral words, (criterion: *r(39) =0.324, p<0.05*; d′: *r(39) =0.336, p<0.05*), showing that higher scores in this category are associated with an increased bias and sensitivity for neutral words.

Distress score: Again no correlations reached our criteria of p<0.01. The analyses revealed a weaker correlation between the distress rating score of the 'Sensory Experience from an Unexplained Source' subscale and both the criterion and d′ for the negative words conditions however (criterion: *r(39) =0.328, p<0.05*; d′: *r(39) =0.329, p<0.05)*. This showed an increased bias and sensitivity for negative words, when participants scored higher on this category.

Distraction Score: No correlations were significant at the level of $p<0.01$. The analyses revealed a weaker correlation between distraction ratings on the 'Changes in Sensory Intensity' subscale, and the criterion and d′ for the neutral word conditions however (criterion: $r(39) = 0.358$, $p<0.05$; d′: $r(39) = 0.362$, $p<0.05$). This revealed that higher distraction rating scorers on this subscale had a decreased bias and increased sensitivity to neutral words. In addition there was a significant correlation between this subscale and the criterion for negative words, $r(39) = 0.351$, $p<0.05$, showing a decreased bias towards negative words for participants who scored high on that subscale.

Frequency Score: Two correlations met the $p<0.01$ criteria. The 'Sensory Experience from an Unexplained Source' subscale correlated with both the criterion and d′ scores for negative words (criterion: $r(39)=0.419$, $p<0.01$; d′: $r(39)=0.432$, $p<0.01$). Participants who rated their experiences on this subscale as more frequent, were also more sensitive to (and less biased towards) negative words. There was also a weaker correlation between the 'Verbal Hallucinations' subscale and the criterion for the positive word conditions, $r(39) = 0.323$, $p<0.05$. Participants who rated their experiences as more frequent were less biased towards positive words.

Summary

ANOVA analyses of the sound detection task showed that emotional words are associated with increased bias, sensitivity and confidence in presence than non-emotional words. This is particularly evident in the cue conditions. These variations in sound detection performance did not differ as a function of hallucination proneness however. Correlational analysis also showed no association between hallucination proneness and sound detection measures, but did show stronger correlations between hits and false positives for low hallucination prone participants. Also the high hallucination prone participants showed a borderline association between vividness and criterion, and between confidence in false alarms and valence ratings.

Finally correlations were calculated between CAPS category scores and signal detection measures for positive, negative and neutral words. This revealed associations between performance for negative words and higher scores on the ‘Sensory experiences from an unexplained source’ subscale. Performance for neutral words was weakly associated with higher scores on the ‘Changes in Sensory Intensity’. Criterion for positive words was weakly associated with increased frequency scores for the ‘Verbal Hallucinations’ subscale.

Discussion

The current auditory word detection task revealed a general bias and sensitivity to reporting the presence of emotional compared to neutral words in noise. This effect was strongest for negative words, followed by positive words, though the two emotional categories did not differ significantly from each other. In addition there was a trend towards greater confidence in the presence of negative words compared to neutral and positive words, showing that confidence ratings tend to follow the pattern found in the criterion and d′ data.

Cue presence affected both the bias and the sensitivity to the target, so that participants were actually more sensitive to the target in the no cue condition. This therefore suggests that presence of the cue causes people to confuse their imagery for the target, resulting in more false positive responses to items in this condition, and also decreased sensitivity. Support for this theory is provided by the finding that despite increased sensitivity to the target in the no cue condition, participants were actually more confident (incorrectly) of hearing the words in the cue condition.

Previous studies revealed that imagining an auditory target can influence sound detection, though the direction of the effect has differed across studies. For instance Segal and Fusella (1970) found an interfering effect whereas Farah and Smith (1983) found a facilitating effect. In the current study provision of a detection cue (i.e. the spoken word which participants imagined and listened for in white noise) was compared to a no cue condition to manipulate the involvement of imagery. Overall the data support Segal and Fusella's (1970) finding that imagining a target sound impairs its subsequent detection. In addition, the cue most affected detection of negative words, followed by positive words, but the cue presence did not affect sensitivity to neutral words. The data suggest that imaging emotional words impairs detection of such words in noise.

The current study also found no main effect of hallucination proneness on detection, nor did this factor interact with either emotion or condition. Correlations between hits and confidence ratings revealed differences between high and low hallucination-prone participants however. The high hallucination-prone group had a lower correlation between their hit rate and confidence rating in both conditions compared with the low hallucination prone group. Mintz and Alpert (1972) reported a similar result suggesting that a lower correlation between hits and confidence ratings was indicative of a reality monitoring disturbance, as this indicated an impaired ability to determine accuracy of an internal state. The current finding suggests that the high hallucination prone participants were less accurate at judging their own internal state, and, in individuals not prone to hallucinations, this effect was modulated by increasing the ability to use auditory imagery (i.e. in the cue condition). Chapter 5 found the opposite result as high hallucination prone participants had a stronger correlation between hit rate and confidence rating. The difference between the two studies may be a function of the sound category used. Verbal stimuli may influence the reality monitoring abilities of high hallucination prone participants more (as found in Chapter 6), whereas these abilities remain relatively unaffected with other sound stimuli (such as in Chapter 5).

Measures of hallucination-proneness

One criticism of the present research is that the LSHS-R score primarily determined the distinction between high and low hallucination-prone individuals. The LSHS-R requires participants to rate the extent to which they have had experiences that range from common mental events (i.e. ‘No matter how hard I try to concentrate, unrelated thoughts always creep into my mind’) to more extreme hallucination experiences (i.e., ‘I have been troubled by hearing voices in my head’). As noted in Chapter 5, an issue here is that the questionnaire cannot distinguish between a person who has had one isolated hallucination-like experience and a person who regularly has such experiences. In such cases the latter individual is clearly more ‘hallucination-prone’ than the former, but they may score the same on the questionnaire. To mitigate this problem a subset of the current participants completed the CAPS (Bell et al., 2006). This relatively new questionnaire is similar to the LSHS-R in that participants answer questions about vivid mental events and hallucination-like experiences, but it also covers other anomalous experiences in different sensory modalities. Participants also rated the frequency, distraction and distress associated with each item they had experienced. This may provide a much more accurate indication of hallucination-proneness. In addition, the questionnaire contained subsets of questions associated with experiences related to hallucination, such as sensory monitoring disturbances and changes in sensory intensity.

Correlations with the rating scores for the subscale questions revealed that scores on ‘Sensory Experience from an Unexplained Source’, ‘Changes in Sensory Intensity’ and ‘Verbal Hallucination’ questions weakly correlated with different measures. In particular, participants who were most sensitive to negative words tended to score higher on measures of distraction on the ‘Changes in Sensory Intensity’ subscale and on distress and frequency measures on the ‘Sensory Monitoring’ subscale. Increased sensitivity to positive words was associated with higher frequency ratings on the ‘Verbal Hallucination’ subscale, while participants who were more sensitive to neutral words generally had higher distraction ratings on the ‘Changes in Sensory Intensity’ subscale. These findings raise various issues. First they suggest an association between different patterns of symptoms and sensitivity to positive words compared to the pattern for negative words. Participants who experience greater distress by sensory experiences from unexplained sources, and who frequently have such experiences, may focus more on negative information because of the distress they experience from them. This may in turn lead to increased sensitivity to negative words. In contrast participants who are more sensitive to positive words scored higher on the ‘Verbal hallucinations’ subscale. This may indicate that such participants have verbal hallucination- like experience in everyday life, but unlike psychotic individuals, do not have negative symptoms such as depression or low self-esteem, which contribute to negative hallucination development (i.e. derogatory voices). This may result in participants being more likely to focus on positive rather than negative information. Further investigation with the use of personality questionnaires would be beneficial in future studies to determine whether participants with an increased negative symptoms are more sensitive to negative words.

Effects of word attributes

In addition to the main factors investigated in this experiment, the study also investigated a number of normative variables associated with the words for their possible influence on signal detection performance. These included: the arousal rating associated with the words (i.e. the degree of stimulation evoked by the word), the valence of the word (i.e. the degree of attraction or aversion evoked by the word) and the auditory imagery vividness of the word. Word arousal correlated with the imagery vividness rating, revealing an association between words with a higher arousal rating and increased auditory imagery vividness. This supports Kensinger and Corkin (2003) who proposed that increased arousal can increase memory for stimuli. Consistent with this, controlling for the effect of arousal (through covariate analysis) eliminated the effect of emotion on imagery vividness and memory, suggesting that the differences between these word categories was due to increased arousal associated with emotional words.

The vividness of auditory imagery for the words also seemed to contribute to response bias. The analysis revealed a trend towards a lower criterion (i.e. increased bias) for higher vividness sounds, suggesting that the higher the vividness the more likely participants were to believe the sound was present. Apparently the vividness of the auditory image can influence the willingness of participants to believe in a word′s presence. Use of more trials may have increased this effect.

In addition, the study revealed an association between valence rating of the words and confidence ratings of high hallucination-prone participants when making false positive responses. The more emotionally positive the target word was, the more confident these individuals were that the words were present. Generally, research suggests that schizophrenic patients tend to have quite emotive hallucinations (Nayani et al., 1996). In the current study, false positives responses were considered as analogous to hallucination-like experiences (i.e. misattribution of internal events as external – Bentall & Slade, 1985) leading to the hypothesis that high hallucination prone participants are more confident in hearing lower valence words (i.e. negative words). This theory is based on investigation of psychiatric populations however, and recent research has suggested that non-psychiatric voice hearers tend to experience predominantly positive hallucinations (Honig, Romme, Ensink, Escher, Pennings & Devries, 1998; Andrew, Gray & Snowden, 2008). The current participants were undergraduate students, rather than psychiatric patients, which may explain why the high hallucination prone participants were more confident about hearing the positive words, rather than the negative words. Further research with a psychiatric sample is needed to test these proposals (Nayani et al., 1996).

General Discussion

The current study investigated how the emotional connotation of a word influences auditory imagery vividness and memory, and whether imagery vividness and valence interact to affect sound detection performance. We also investigated how the emotional content of the stimuli affected auditory imagery and detection tasks of hallucination prone participants. Previous studies have found that hallucination-prone participants make more reality monitoring errors to emotional words (Laroi, Marczewski, Van der Linden, & Evrard, 2003). The current study revealed that emotional words received higher in auditory imagery vividness ratings, are remembered more than neutral words and are also detected easier in noise. These effects were independent of hallucination proneness overall however. Hallucination proneness did affect memory, showing that high hallucination-prone participants remembered more words than low hallucination-prone participants. This is against what might be predicted from research into memory abilities of patients with schizophrenia, as memory impairments are a frequently reported deficit in this group (see Aleman, Hijman, de Haan, & Kahn, 1999, for a review). In the current experiment however, two thirds of the stimuli were emotional in content and we proposed that high hallucination-prone participants had better memory due to the increased arousal associated with such words. Indeed, controlling for arousal rating removed the effect of hallucination-proneness, suggesting that differences in arousal between words differentially affected the memory of high and low hallucination-prone participants.

Correlation analysis revealed a weaker association between hits and confidence ratings for high compared to low hallucination prone participants. This suggested that these participants may have some reality monitoring disturbance in comparison to low hallucination-prone participants (Mintz et al., 1972). This is in contrast to the previous sound detection study in Chapter 5. This analysis revealed no difference in the correlation between hits and confidence ratings, for the high and low hallucination prone participants in the cue condition. This correlation was stronger in the no cue condition however. This may be due to the difference in stimuli between the two studies. The previous study incorporated all categories of sound, while the current study focused on language. Therefore the poorer reality monitoring performance may be specific to language.

A possible reason for the lack of strong effects of hallucination proneness in both experiments was that the median values defined the high and low hallucination proneness groups, rather than the extreme scores. Analysis of extreme scorers still resulted in no differences between the two groups however. Another factor that may explain the lack of differences is the design of the hallucination proneness questionnaire, which does not taken into account frequency, distraction or distress caused by hallucination-like experience. Investigation of these factors in relation to imagery vividness and sound detection of emotional material may reveal a more detailed picture of the relationships between hallucination proneness, emotional content, vividness and sound detection. Indeed data from a subset of participants who completed the CAPS questionnaire, revealed an association between performance in negative word conditions and source monitoring and sensory intensity disturbances. Further studies could investigate how personality characteristics interact with emotional word detection.

In conclusion the current study revealed an association between words higher in emotional content or arousal and increased auditory imagery vividness, and increased recall. In addition there was better detection of words with greater emotional content suggesting that arousal associated with the word can increase sensitivity to it in noise. There was little effect of hallucination proneness, however, suggesting that the effect of emotion on word detection may reflect a general trait which is unrelated to hallucination-like experiences.

Chapter Seven. General Discussion

The present thesis provides an exploratory analysis of factors that contribute to the experience of auditory imagery vividness. The primary aim was to assess (i) how factors associated with particular sounds (i.e. category, familiarity and emotion) and (ii) how the presence of imagery cueing affect subjective experiences of auditory imagery vividness. A second aim was to evaluate how these factors affect the interaction between imagery and perception. A third aim was to investigate how proneness to hallucinations influences both subjective experience of imagery vividness, and the interaction between imagery and perception. This discussion focuses on the outcome of these aims in turn.

1. Factors that influence imagery vividness

A key aim of this thesis was to investigate how cognitive factors affect auditory imagery vividness. A limited amount of research has investigated this, as the majority of auditory imagery research focused on the similarities between imagery and perception in terms of acoustic characteristics (Halpern et al., 2004; Intons-Peterson, 1980; Intons-Peterson, Russell, & Dressel, 1992) and neural activation (Shergill et al., 2001; Yoo et al., 2001; Bunzeck et al., 2005).

In Chapter 2, Experiment 1 and 2 investigated how sound category and familiarity influenced vividness of auditory imagery. These studies showed that participants consistently gave higher ratings to music and speech sounds, compared to animal and environmental sounds. Possibly music and speech items can utilise subvocalisation to enhance imagery vividness; in contrast animal and environmental sounds may be more difficult to subvocalise and hence give rise to weaker imagery ratings. Also in our everyday lives, we may experience imagery for music (Bailes, 2006) and speech more often than that for animal and environmental sounds. Sound familiarity highly influenced vividness of imagery, with unfamiliar sounds in all sound categories receiving lower ratings than familiar sounds. This effect was strongest for animal and environmental sounds, probably because we have more experience of unfamiliar music and speech, compared to animal and environmental sounds. Baddeley and Andrade (2000) also investigated how sound familiarity affects auditory imagery vividness, finding that vividness ratings were lower for unfamiliar items. The current study therefore supports Baddeley and Andrade's (2000) finding and expands on it to show that sound category can also influence imagery vividness.

It could also be suggested however, that variation in vividness ratings according to sound category may be a function of the familiarity associated with the sounds. To assess this, Experiment 3a obtained vividness and familiarity ratings for sounds in different categories, and assessed correlations between these ratings. Though vividness and familiarity ratings showed good correlation for all sound categories, the study showed that familiarity could not solely explain vividness variations across sound categories. The study still found higher vividness ratings for music and speech items, compared to animal and environmental sounds, even with familiarity rating as a covariate in the analyses. This shows that vividness ratings are not purely a function of familiarity, and that other factors must contribute to the experience of vividness.

In addition this experiment also revealed that auditory imagery ratings are reliable across time, as test-retest analysis revealed that there was no difference in vividness ratings obtained a week apart from each other.

Experiment 1 also investigated cross-modal interactions between auditory and visual imagery for different sound categories. Here, it was hypothesised that a picture cue to imagine the target sound would increase auditory imagery vividness, as the cross-modal association would be stronger between these pairings, than a verbal imagery cue. This hypothesis was supported: picture cues resulted in higher imagery vividness than name cues. This extends Lehmann and Murray (2005) finding of improved memory recall for items encoded as cross-modal memories (e.g. auditory-visual) to suggest that cross-modal associations also increase the vividness associated with a memory. The study also revealed particularly high imagery vividness for animal and environmental sounds cued with picture cues, suggesting a stronger association between the visual and auditory representation of these items. This may be because sounds of these items are most commonly experienced with a visual presentation of the item. Further research should focus on whether the same holds for visual imagery vividness to determine whether the cross-modal association is as strong for this modality also.

Experiment 2 expanded on Baddeley and Andrade's findings by investigating how hearing either white noise, or a sound from the same or different sound category as the target, affected imagery vividness for that sound. This study hypothesised that vivid images are more perceptual-like in nature (Aleman et al., 2003) and that imagery and perception activate similar neural regions (Yoo et al., 2001). Therefore the study hypothesised that listening to a sound while attempting to imagine the target sound would disrupt the imagery process, as fewer resources would be available for sound imagery. It was also hypothesised that this would disruption would increase as the image and interference sound became more semantically similar (i.e. from the same sound category). The study revealed that, although listening to meaningful distractor sounds reduced auditory imagery vividness compared to white noise, listening to distractors from the same versus different sound category resulted in lower ratings for music targets only. Further investigation revealed that participants rated music targets and interference sounds as more similar to each other, compared to same-category pairings of other sound categories. This congruency effect found for music target sounds is likely to be due to acoustic similarity between the sounds; high acoustic similarity increases the difficulty in distinguishing the target image from any distractors, and therefore lowers imagery vividness ratings for these pairings. This is contrary to the hypothesis that semantic similarity would result in lower imagery vividness, and suggests that acoustic similarity between images and sounds can affect imagery vividness.

A criticism of vividness ratings is that they are largely a subjective opinion of the participants own imagery abilities. As such, “vividness” cannot be observed directly and we cannot be sure if one person’s vivid image is the same as another’s. In addition vividness may be biased by perceived socially desirable responses (Allbutt et al., 2008). In effort to provided objective evidence of the validity of vividness ratings, Chapter 3 determined the neural correlates of vividness. This study investigated the association between imagery vividness ratings for animal and environmental sounds, and activation strength during imagery. Cui, Jeter, Yang, Montague, and Eagleman’s (2007) study of visual imagery revealed that participants who had high imagery vividness in a visual imagery questionnaire also had higher activation strength in early visual areas. The current fMRI study went one step further, by asking participants to rate their imagery vividness for sounds during the fMRI experiment. The areas modulated by imagery vividness ratings revealed significant activation in the left MFG and bilateral insula which previous studies found active during auditory imagery, inner speech, sound retrieval and auditory attention (Aleman et al., 2005; Hoshiyama et al., 2001; Shergill et al., 2001; Voisin et al., 2006; Yoo et al., 2001). Therefore this study suggests imagery vividness ratings do reflect neural activation in areas responsive to auditory imagery. This activation pattern, and evidence from the studies reported above, suggests the involvement of the MFG and insular regions in judgement of one's own auditory imagery vividness.

Finally Experiment 7a assessed whether imagery vividness and recall memory would be affected by the emotional content of the imagined item. Participants rated positive and negative words as more vivid than neutral words, and recalled more of the emotional words than neutral. In addition the significant association between vividness ratings and arousal ratings suggests that valenced words may increase the participants' arousal levels and leading to increases in imagery vividness. Thus, as well as the type of sound imagined, the arousal associated with the sound can also affect subjective imagery ratings.

These studies reveal that imagery vividness is stable and robust. Vividness rating can be affected by factors associated with the sound, such as category, familiarity and arousal, and also by task-related factors, such as cues to imagine, and dual task performance. Finally the fMRI study revealed activation in frontal regions of the brain that was associated with greater imagery vividness. Together these studies suggest that imagery vividness ratings may not be as idiosyncratic and unreliable as first thought. Further evidence for this is discussed below.

2. Interaction between imagery and perception

The main criticism of using vividness ratings as a measure of auditory imagery is that they are subjective and may be prone to idiosyncratic construction and comparison to the perceived vividness of others imagery (Aleman et al., 1999; Allbutt et al., 2008). Because of such criticisms further experiments aimed to establish more objective evidence of the influence of vividness ratings. Experiment 3b determined whether the variations in vividness observed revealed in Experiments 1, 2 and 3a were indicative of actual imagery strength. A sound detection task was used because previous studies found that imagining a sound can have a facilitatory or inhibitory effect on detecting it in noise, depending on how closely the image and target sounds match (Farah et al., 1983; Segal et al., 1970). The current study hypothesised that, because provision of a valid cue would trigger a matching auditory image, detection would be facilitated. This hypothesis was partially confirmed as cue provision resulted in a greater bias toward believing a sound was present, as well as greater confidence in the target's presence. This offers some support for (Farah et al., 1983) findings, as cue presence increased bias and confidence in the targets presence.

In addition the study assessed differences in detection between high and low vividness sounds, and familiar and unfamiliar sounds. This showed that vividness only influenced participants' bias and confidence in a sound's presence (both these measures were greater for higher vividness sounds) but sensitivity was not affected. Familiarity also did not alter sensitivity, but did influence the bias, but in interaction with vividness.

For low vividness sounds, peak bias and confidence occurred when the sounds were familiar, whereas for high vividness sounds, the peak occurred when the sounds were unfamiliar. This interaction can be explained by differing levels of cognitive effort involved in detecting high and low vividness sounds that vary in familiarity. High vividness, familiar sounds are easy to imagine, therefore less cognitive effort is required to detect them correctly in noise. Unfamiliar images however, have a less well-defined 'template' and so are more difficult to distinguish, and require more cognitive resources for detection, when presented in noise. Segal and Fusella's study provided evidence for this, as imagining an unfamiliar sound prior to detection impaired tone detection more that imagining a familiar sound. Low vividness sounds may also require more cognitive effort to generate the image, especially when the sound is unfamiliar. In this case image generation is difficult, so participants are more conservative about judging when their images match the target.

The fMRI study in the current thesis directly investigated the interaction between imagery and perception, by determining whether they shared neural resources. Though the study found no overall activation for the imagery conditions when compared to baseline, it did find greater activation in the right STG to animal sound imagery compared to environmental sound imagery. This finding is similar to that found for the same contrast between animal and environmental sound perception. This study therefore supports previous studies findings that imagery and perception utilise a similar neural network (Bunzeck et al., 2005; Halpern et al., 2004; Shergill et al., 2001; Yoo et al., 2001).

Contrasts in the opposite direction (i.e. environmental sound perception/imagery vs. animal sound perception/ imagery) revealed no significant activations. These findings mirror those found by Kraut et al. (2006), who reported right STG activation for contrasts between animal sound perception/imagery and environmental sound perception/imagery, but no differential activation for the opposite contrast. Therefore animal sounds may form a more distinct semantic group that is more strongly associated with sound memory than environmental sounds.

An additional analysis determined connectivity between activity in response to imagery (i.e. right STG) and other regions of the brain, to assess how these regions are related to imagery production. This PPI analysis revealed frontal activation when imagery conditions were contrasted with perception. This suggests the involvement of frontal region in auditory image generation, and that long-term memory for sounds is fed forward from the right STG to these frontal regions. This may either support image representations for the stimuli or judgements about the stimuli for the imagery task.

Appendix I examined the role of the frontal regions in auditory processing through a case study of a patient with left frontal lobe damage. The study investigated apperceptive and associative deficits for different sound categories and determined whether similar deficits occurred across modalities also (e.g. in vision), . DS showed impairments in both auditory and visual versions of the tasks, although more so for the auditory versions. The greatest impairment was for musical instruments, followed by animal sounds and environmental sounds, and in all cases there was a mixture of apperceptive and associative symptoms. Previous research with patients with temporal lobe impairments has found that patients with left lesions have more associative impairments while patients with right lesions have apperceptive impairments (Vignolo, 1982). The current study suggests that frontal lesions may lead to general recognition impairments covering both forms of deficit, rather than to either ‘pure’ perceptive or semantic impairments. Auditory recognition abilities of patients with frontal lesions have not been investigated before. The fMRI study in Chapter 3 however, found frontal activation during perception and imagery for sounds and a significant coupling between temporal and frontal lobe regions during imagery. Therefore it is possible that frontal regions play a role in sound recognition, mediated by connectivity to the temporal lobes, which would explain why DS displayed auditory recognition deficits. These neuropsychological data indicate involvement of the frontal lobes in producing imagery in the first place, as well as in the evaluation of such imagery.

3. Hallucination proneness

One area where the interaction between imagery and perception is most poignant is in the phenomena of auditory hallucination. Knowledge about the specific link between imagery and hallucinations is lacking, though it is generally agreed that they involve the miscategorisation of auditory images as coming from an external source. The theory that hallucination prone participants have different levels of imagery vividness compared to non-hallucination prone participants is not new, but research into this area has produced some mixed findings (Aleman et al., 2001; Brett et al., 1977; Mintz et al., 1972; Sack et al., 2005; Starker et al., 1982).

Chapter 4 investigated the relationship between hallucination proneness, schizotypy, anomalous experiences, and imagery vividness. Analysis revealed good association between different schizotypy measures, which is likely to be due to these measures containing very similar items. This study revealed no association between these measures and auditory imagery vividness however. This supports Starker and Jolins's (1982) findings of no imagery vividness differences between high and low hallucination prone participants. The study also investigated types of anomalous experiences that may be related to imagery vividness. The CAPS questions about unusual experiences in all modalities, but it also has subscales based on particular types of anomalous experiences. Stronger auditory imagery vividness was associated with the sensory flooding subscale. This suggests that vivid images have a more perceptual-like quality, and that people who experience such imagery tend to experience over stimulation from sensations.

Other research suggests that it is not vividness of imagery per se that differs between high and low hallucination prone participants, but rather the attribution of auditory image (Bentall et al., 1985). Chapter 5 also employed the same sound detection paradigm as Chapter 2, but with the extra variable of hallucination proneness included. Here we anticipated that, since hallucination prone participants have difficulty in correctly attributing an image to internal sources, they would make more false positive errors in the sound detection task, particularly when the sound was highly vivid. Hallucination proneness did not affect the response bias, sensitivity or confidence in responses however, suggesting that hallucination proneness does not influence detection of sounds in noise. In addition correlation analysis assessed reality monitoring abilities, to see whether high and low hallucination prone participants differ in this ability. Mintz and Alpert (1972) stated that stronger correlations indicate that participants are more accurate about their internal state. In the current study, this analysis revealed a stronger correlation for high hallucination prone participants in the no cue condition. This is contrary to the hypothesis that the high hallucination prone group would be more impaired at judging their internal state. This suggests that in this condition, high hallucination prone participants were better at judging their internal state when they did not received any extra information about the target. This implicates that distraction from external information may impair reality monitoring abilities of this group of participants.

In Chapter 5 the sounds were emotionally neutral. Given that hallucinatory experiences are often emotionally charged, the next chapter investigated vividness and detection abilities of high and low hallucination prone for emotional compared to neutral words. Experiment 7a assessed whether imagery vividness and recall memory varied with the emotional content of the target words. In addition the study predicted that this effect would be most pronounced in high hallucination prone participants. Though positive and negative words were rated as more vivid than neutral words, hallucination proneness did not affect the imagery ratings, nor did it interact with the emotion of the word. High hallucination prone participants did recall more words in this experiment, though again, emotional connotation did not interact with this. Therefore this study showed that contrary to the hypothesis, the affect of emotion on vividness and memory recall is not specific to hallucination proneness.

Experiment 7b again employed the sound detection paradigm, this time to determine whether high and low hallucination prone participant differ in their ability to detect emotional and neutral words in noise. There was an association between increased response bias, confidence and decreased sensitivity, for emotional words compared to neutral words, particularly when participants were cued with which item to expect. There was no effect of hallucination proneness however. Again this suggests that the effects of emotional connotation are not specific to hallucination prone participants. Experiment 7a also showed an association between auditory imagery vividness ratings and arousal. It was suggested that arousal could increase the vividness associated with the words making any image more ‘perceptual-like’. This in turn may make it more difficult to determine whether the image comes from an internal or external source. Experiment 7b also revealed reduced sensitivity in the cue condition, which was unusual, since participants received cues for which word to listen for. A likely explanation for this is that the bias created by the emotional words drags down perceptual sensitivity to the words, in the cue condition. In contrast there was no difference between the cue and no cue conditions for neutral words in either the bias or in sensitivity. This provides evidence for the fact that having an emotional connation associated with words inhibits detection of it in noise.

Chapter 6 also assessed reality monitoring abilities by correlating hits and confidence ratings. This correlation was lower for high hallucination prone participants in both conditions, suggesting that this group were more impaired in reality monitoring than the low hallucination prone group. This is in contrast to the same analysis in Chapter 5. It follows that emotional, verbal material may impair reality monitoring abilities (Chapter 6) more than neutral material (Chapter 5), for high hallucination prone participants.

In conclusion, these studies showed little evidence for an association between hallucination proneness and auditory imagery vividness. Hallucination proneness also had little effect on sound detection, whether the material contained non-verbal neutral sounds, or emotional verbal sounds. In the latter case hallucination proneness was associated with impaired reality monitoring abilities, suggesting that emotional material can disrupt such participant's accuracy about their internal state.

4. Strengths and Limitations

The following section highlights the strengths and limitations of the key components of this thesis.

4.1. Vividness Ratings

Strengths: Vividness of auditory imagery has received little previous research, with the majority of studies focusing on the use of vividness questionnaires and their association with more objective measures of imagery. Only one previous study investigated how different cognitive factors effect vividness, such as familiarity and dual tasks (Baddeley et al., 2000). Therefore the current thesis provides a more detailed investigation of auditory imagery vividness than previous studies, investigating the roles of sound category, familiarity, cross-modal cueing and hallucination proneness. Previous studies suggested that imagery can affect the detection of a faintly presented sound in noise (Segal et al., 1970; Farah et al., 1983) and that vivid images have a more perceptual-like quality (Aleman et al., 2003). No studies have investigated how auditory imagery associated with a sound influences sound detection however, (though Segal et al., 1970, noted that imagining an unfamiliar compared to familiar sound impairs detection).

Limitations: The present auditory imagery studies investigated the effect of cognitive factors on imagery vividness and to assess the effect of these factors on sound detection. Use of these measures determined the relationship between subjective and objective measures of imagery vividness. Though the validity of vividness ratings was established, it would have been beneficial to see how vividness ratings relate to other objective measures of auditory imagery vividness apart from sound detection.

4.2. Sound Detection

Strengths: The use of real sounds in the sound detection paradigm here is also fairly novel, as the majority of studies use single words or tones as stimuli (Bentall et al., 1985; Barkus et al., 2007; Sack et al., 2005; Aleman et al., 2001). An advantage of the current stimuli is that they are more cognitively engaging stimuli and they may better mimic hallucination-like experiences in the real world.

Limitations: Employment of such a wide range of stimuli also generated problems, as contrasting sound categories differed from each other in their acoustic characteristics, such as HNR and RMS. Therefore the RMS was adjusted to ensure equal detectability of different sounds at threshold. Covariate analyses showed that that HNR (but not RMS) contributed to some but not all of the significant effects (see Appendix H for these further analyses).

4.3. Hallucination Proneness

Limitations: The studies reported in this thesis revealed little effect of hallucination proneness on sound detection performance compared to previous studies (Barkus et al., 2007; Bentall and Slade, 1985). The study grouped participants by the median scores on the hallucination proneness questionnaire, rather than by extreme values, which may have formed more distinct groups. Chapter 4 did however prove that the sample of participants in the current studies is similar to that of previous studies, and some previous studies have found differences between high and low hallucination prone participants in performance on signal detection tasks, using a median split (Aleman et al., 1999). Despite this, there was little effect of hallucination proneness even when analyses focussed on the top and bottom 25th percentile scorers only. This may be because there truly was little difference in sound detection performance between high and low hallucination prone participants.

On the other hand, previous studies question the validity of dichotomising continuous independent variables (Maxwell & Delaney, 1993; MacCallum, Zhang, Preacher, & Rucker, 2002). These studies claim that dichotomising such measures results in the loss of a large amount of information through data reduction, which can lead to loss of power and increased chances of finding spurious effects. Therefore there is a case for keeping independent measures continuous in statistical analyses, through the use of regression analyses

Regression analysis was not appropriate for analysis of the data from the sound detection tasks in the current thesis, as the aim was to determine the interaction between the experimental manipulations (i.e. vividness, familiarity, cueing, emotion) and hallucination proneness. Correlations between hallucination proneness scores and the sound detection measures were assessed however, which yielded no significant effects. Future investigation into the influence of hallucination proneness on sound detection performance should endeavour to maintain the continuity of hallucination proneness.

Strengths: Though hallucination proneness appeared to have fairly limited effects on the signal detection tasks in the current studies, a number of different measures of hallucination proneness were assessed (particularly in Chapter 4). This enabled a clearer picture of the components of hallucination proneness and the relationship between different measures of hallucination proneness, anomalous experience and positive schizotypy.

4.4. Sparse sampling auditory fMRI

Strengths: Like other previous investigations of auditory imagery/perception, Chapter 3 employed a sparse sampling approach. The particular technique used here differed from those in previous studies as it involved an event related sequence with a variable scan time of between 4 and 8 sec following a silent period in the scanner. This enabled stimuli to be presented in silence, and so reduce the distracting and masking effects of scanner noise on activation.

Limitations: One issue that arose from this study is the need for a suitable baseline task, which mimics the task demands of the experimental conditions, but also inhibits the use of verbal imagery or inner speech. The current study found negative beta values in the imagery conditions, as it is likely that participants used inner speech during the baseline condition. Because auditory imagery and inner speech involves similar brain regions (Shergill et al., 2001) it is hypothesised that contrasting the imagery condition with the baseline removed the activation in the imagery conditions.

Also the current study measured auditory imagery vividness during the actual fMRI task, which enabled investigation of areas of the brain particularly responsive to variations in imagery vividness. This is important as it provides evidence that imagery vividness ratings have an actual physiological basis, rather than being purely based on socially desirable responding or individual idiosyncrasies.

5. Further research

This thesis has revealed that vividness ratings are stable, that a number of factors can affect them and that imagery vividness can influence detection of sounds in noise. In addition the thesis revealed some of the neural underpinnings of vividness ratings. Further research with vividness ratings could investigate their association with subvocalisation, and how preventing subvocalisation influences vividness ratings. Such studies would reveal details about the creation and manipulation of imagery vividness of imagery. An adequate baseline that prevents auditory imagery use would improve the current fMRI study design. Such a study could reveal neural activation in association to unfamiliar, non-verbal sounds, to determine how their neural activation differs from that of familiar sounds. This may also reveal why vividness ratings are lower for unfamiliar compared to familiar sounds. In addition it would be interesting to image the neural activation in response to false positive responses in the sound detection task (Barkus et al., 2007). This would determine whether auditory imagery does play a role in sound detection.

Investigation of the effects of vividness, familiarity and emotional connotation on sound detection performance of hallucinating and non-hallucinating patients would be beneficial. This would determine whether these factors do play a role in actual hallucination experiences, and to determine how such participants differ from hallucination prone control participants.

6. Conclusions

In conclusion, the current thesis examined auditory imagery, and found that cognitive factors such as sound category, familiarity, cueing and sound perception affect people's subjective experiences of auditory imagery vividness. The rated vividness of a sound interacted with its familiarity to influence the detection of sounds in noise. The sparse-sampling fMRI study revealed that auditory imagery and perception for non-verbal sounds overlap in activation in areas associated with auditory processing. In addition this study determined regions responsive to auditory imagery vividness ratings, revealing a role of the MFG and bilateral insular in the judgement of one's own auditory imagery. The second part of this thesis assessed the interactions between hallucination proneness and auditory imagery. Hallucination proneness had little effect on sound detection performance. Correlations between hit rate and confidence rating assessed reality monitoring abilities related to hallucination proneness, showing little difference in correlation strength between high and low hallucination prone participants over all categories of sound. Poorer reality monitoring performance was found for the high hallucination prone group when stimuli were verbal and emotional in content however. This suggests that such material is particularly vulnerable to reality monitoring errors in such participants, possible due to the increased arousal that such stimuli produce. This therefore provides a basis for why patient with schizophrenia commonly experience verbal, emotional hallucinations.

Further research is now need to further determine the organisation of auditory processing areas in response to auditory imagery vividness, and how vividness can interact with perception of sounds, particularly in hallucinating and non-hallucinating patients.

References

Aleman, A., Bocker, K. B. E., & de Haan, E. H. F. (1999). Disposition towards hallucination and subjective versus objective vividness of imagery in normal subjects. *Personality and Individual Differences, 27,* 707-714.

Aleman, A., Bocker, K. B. E., & de Haan, E. H. F. (2001). Hallucinatory predisposition and vividness of auditory imagery: Self-report and behavioral indices. *Perceptual and Motor Skills, 93,* 268-274.

Aleman, A., Bocker, K. B. E., Hijman, R., de Haan, E. H. F., & Kahn, R. S. (2003). Cognitive basis of hallucinations in schizophrenia: role of top-down information processing. *Schizophrenia Research, 64,* 175-185.

Aleman, A. & de Haan, E. H. F. (2004). Fantasy proneness, mental imagery and reality monitoring. *Personality and Individual Differences, 36,* 1747-1754.

Aleman, A., Formisano, E., Koppenhagen, H., Hagoort, P., de Haan, E. H. F., & Kahn, R. S. (2005). The functional neuroanatomy of metrical stress evaluation of perceived and imagined spoken words. *Cerebral Cortex, 15,* 221-228.

Aleman, A., Hijman, R., de Haan, E. H. F., & Kahn, R. S. (1999). Memory impairment in schizophrenia: A meta-analysis. *American Journal of Psychiatry, 156,* 1358-1366.

Aleman, A., Nieuwenstein, M. R., Bocker, K. B. E., & de Haan, E. H. F. (2000). Mental imagery and perception in hallucination-prone individuals. *Journal of Nervous and Mental Disease, 188,* 830-836.

Allardyce, J., Gaebel, W., Zielasek, J., & van Os, J. (2007). Deconstructing Psychosis Conference February 2006: The validity of schizophrenia and alternative approaches to the classification of psychosis. *Schizophrenia Bulletin, 33,* 863-867.

Allbutt, J., Ling, J., Heffernan, T. M., & Shafiullah, M. (2008). Self-Report Imagery Questionnaire Scores and Subtypes of Social-Desirable Responding Auditory Imagery, Visual Imagery, and Thinking Style. *Journal of Individual Differences*, *29,* 181-188.

Altmann, C. F., Doehrmann, O., & Kaiser, J. (2007). Selectivity for animal Vocalizations in the human auditory cortex. *Cerebral Cortex, 17,* 2601-2608.

Amaro, E., Williams, S. C. R., Shergill, S. S., Fu, C. H. Y., MacSweeney, M., Picchioni, M. M. et al. (2002). Acoustic noise and functional magnetic resonance imaging: Current strategies and future prospects. *Journal of Magnetic Resonance Imaging, 16,* 497-510.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders (4th ed.).* American Psychiatric Association, Washington, DC.

Andersson, J. L. R., Hutton, C., Ashburner, J., Turner, R., & Friston, K. (2001). Modeling geometric deformations in EPI time series. *Neuroimage, 13,* 903-919.

Andreasen, N. (1984). *The scale for the assessment of positive and negative symptoms (SAPS and SANS).* Iowa City, IA, USA: Iowa University Press.

Andreasen, N. C. (1982a). Negative Symptoms in Schizophrenia - Definition and Reliability. *Archives of General Psychiatry, 39,* 784-788.

Andreasen, N. C. (1982b). Negative V Positive Schizophrenia - Definition and Validation. *Archives of General Psychiatry, 39,* 789-794.

Arndt, S., Alliger, R. J., & Andreasen, N. C. (1991). The Distinction of Positive and Negative Symptoms - the Failure of A 2-Dimensional Model. *British Journal of Psychiatry, 158,* 317-322.

Arnott, S. R., Grady, C. L., Hevenor, S. J., Graham, S., & Alain, C. (2005). The functional organization of auditory working memory as revealed by fMRI. *Journal of Cognitive Neuroscience, 17,* 819-831.

Ashburner & Friston (2003). Image segmentation. In Frackowiak RS, Friston KJ, Frith C, Dolan RJ, Price C, Zeki S, Ashburner J, & Penny W (Eds.), *Human Brain Function* (pp. 635-654). (Academic, New York).

Baddeley, A. (1986). *Working Memory*. Oxford University Press.

Baddeley, A. D. & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology-General, 129,* 126-145.

Baethge, C., Baldessarini, R. J., Freudenthal, K., Streeruwitz, A., Bauer, M., & Bschor, T. (2005). Hallucinations in bipolar disorder: characteristics and comparison to unipolar depression and schizophrenia. *Bipolar Disorders, 7,* 136-145.

Bailes, F. A. (2006). The use of experience-sampling methods to monitor musical imagery in everyday life. *Musicae Scientiae, 10,* 173-190.

Barkus, E., Stirling, J., Hopkins, R., Mckie, S., & Lewis, S. (2007). Cognitive and neural processes in non-clinical auditory hallucinations. *British Journal of Psychiatry, 191,* S76-S81.

Barrett, T. R. (1993). Verbal Hallucinations in Normals .2. Self-Reported Imagery Vividness. *Personality and Individual Differences, 15,* 61-67.

Barrett, T. R. & Etheridge, J. B. (1992). Verbal Hallucinations in Normals .1. People Who Hear Voices. *Applied Cognitive Psychology, 6,* 379-387.

Bearden, C. E., Rosso, I. M., Hollister, J. M., Sanchez, L. E., Hadley, T., & Cannon, T. D. (2000). A prospective cohort study of childhood behavioral deviance and language abnormalities as predictors of adult schizophrenia. *Schizophrenia Bulletin, 26,* 395-410.

Beck, A. T. & Rector, N. A. (2003). A cognitive model of hallucinations. *Cognitive Therapy and Research, 27*, 19-52.

Belardinelli, M. O., Palmiero, M., Sestieri, C., Nardo, D., Di Matteo, R., Londei, A. et al. (2009). An fMRI investigation on image generation in different sensory modalities: The influence of vividness. *Acta Psychologica, 132,* 190-200.

Belin, P., Zatorre, R. J., Hoge, R., Evans, A. C., & Pike, B. (1999). Event-related fMRI of the auditory cortex. *Neuroimage, 10,* 417-429.

Bell, V., Halligan, P. W., & Ellis, H. D. (2006). The Cardiff Anomalous Perceptions Scale (CAPS): A new validated measure of anomalous perceptual experience. *Schizophrenia Bulletin, 32,* 366-377.

Bentaleb, L. A., Beauregard, M., Liddle, P., & Stip, E. (2002). Cerebral activity associated with auditory verbal hallucinations: a functional magnetic resonance imaging case study. *Journal of Psychiatry & Neuroscience, 27,* 110-115.

Bentall, R. (2006). Madness explained: Why we must reject the Kraepelinian paradigm and replace it with a 'complaint-orientated' approach to understanding mental illness. *Medical Hypotheses, 66,* 220-233.

Bentall, R. P. (1992). The Syndromes and Symptoms of Psychosis. In R.P.Bentall (Ed.), *Reconstructing Schizophrenia* (pp. 23-60). Routledge.

Bentall, R. P., Claridge, G. S., & Slade, P. D. (1989). The Multidimensional Nature of Schizotypal Traits - A Factor Analytic Study with Normal Subjects. *British Journal of Clinical Psychology, 28,* 363-375.

Bentall, R. P., Jackson, H. F., & Pilgrim, D. (1988). The Concept of Schizophrenia Is Dead - Long Live the Concept of Schizophrenia. *British Journal of Clinical Psychology, 27,* 329-331.

Bentall, R. P. & Kaney, S. (1989). Content Specific Information-Processing and Persecutory Delusions - An Investigation Using the Emotional Stroop Test. *British Journal of Medical Psychology, 62,* 355-364.

Bentall, R. P., Kinderman, P., & Kaney, S. (1994). The Self, Attributional Processes and Abnormal Beliefs - Towards A Model of Persecutory Delusions. *Behaviour Research and Therapy, 32,* 331-341.

Bentall, R. P. & Slade, P. D. (1985). Reality Testing and Auditory Hallucinations - A Signal-Detection Analysis. *British Journal of Clinical Psychology, 24,* 159-169.

Berger, G. H. & Gaunitz, S. C. B. (1977). Self-Rated Imagery and Vividness of Task Pictures in Relation to Visual Memory. *British Journal of Psychology, 68,* 283-288.

Bijl, R. V., Ravelli, A., & van Zessen, G. (1998). Prevalence of psychiatric disorder in the general population: results of the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social Psychiatry and Psychiatric Epidemiology, 33,* 587-595.

Binder, J. (2000). The new neuroanatomy of speech perception. *Brain, 123,* 2371-2372.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Springer, J. A., Kaufman, J. N. et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex, 10,* 512-528.

Binder, J. R., Frost, J. A., Hammeke, T. A., Rao, S. M., & Cox, R. W. (1996). Function of the left planum temporale in auditory and linguistic processing. *Brain, 119,* 1239-1247.

Birchwood, M. & Chadwick, P. (1997). The omnipotence of voices: testing the validity of a cognitive model. *Psychological Medicine, 27,* 1345-1353.

Birchwood, M. J. & Jackson, C. (2001). *Schizophrenia*. Psychology Press.

Bleuler, E. (1911). *Dementia Praecox or the Group of Schizophrenias*. Zinkin J., trans. New York: International Universities Press.

Boatman, D. F. (2006). Cortical auditory systems: Speech and other complex sounds. *Epilepsy & Behavior, 8,* 494-503.

Bocker, K. B. E., Hijman, R., Kahn, R. S., & de Haan, E. H. F. (2000). Perception, mental imagery and reality discrimination in hallucinating and non-hallucinating schizophrenic patients. *British Journal of Clinical Psychology, 39,* 397-406.

Boyle, M. (1992). The Non-Discovery of Schizophrenia? In R.P.Bentall (Ed.), *Reconstructing Schizophrenia* (pp. 3-22). Routledge.

Bradley, M. M. & Lang, P. J. (1999). *Affective norms for English words (ANEW)* Gainesville, FL. The NIMH Center for the Study of Emotion and Attentions, University of Florida.

Brasic, J. R. (1998). Hallucinations. *Perceptual and Motor Skills, 86,* 851-877.

Brebion, G., Amador, X., David, A., Malaspina, D., Sharif, Z., & Gorman, J. M. (2000). Positive symptomatology and source-monitoring failure in schizophrenia - an analysis of symptom-specific effects. *Psychiatry Research, 95,* 119-131.

Brebion, G., Gorman, J. M., Amador, X., Malaspina, D., & Sharif, Z. (2002). Source monitoring impairments in schizophrenia: characterisation and associations with positive and negative symptomatology. *Psychiatry Research, 112,* 27-39.

Brett, E. A. & Starker, S. (1977). Auditory Imagery and Hallucinations. *Journal of Nervous and Mental Disease, 164,* 394-400.

Brockington, I. F., Kendell, R. E., & Leff, J. P. (1978). Definitions of Schizophrenia - Concordance and Prediction of Outcome. *Psychological Medicine, 8,* 387-398.

Buchtel, H. A. & Stewart, J. D. (1989). Auditory Agnosia - Apperceptive Or Associative Disorder. *Brain and Language, 37,* 12-25.

Bunzeck, N., Wuestenberg, T., Lutz, K., Heinze, H. J., & Jancke, L. (2005). Scanning silence: Mental imagery of complex sounds. *Neuroimage, 26,* 1119-1127.

Caramazza, A. & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate-inanimate distinction. *Journal of Cognitive Neuroscience, 10,* 1-34.

Chadwick, P. & Birchwood, M. (1994). The Omnipotence of Voices - A Cognitive Approach to Auditory Hallucinations. *British Journal of Psychiatry*, *164*, 190-201.

Chapman, L. J., Chapman, J. P., & Raulin, M. L. (1978). Body-Image Aberration in Schizophrenia. *Journal of Abnormal Psychology, 87,* 399-407.

Claridge, G. (1992). Can a Disease Model of Schizophrenia Surive? In R.P.Bentall (Ed.), *Reconstructing Schizophrenia* (pp. 157-183). Routledge.

Claridge, G. (1997). *Schizotypy: Implications for Illness and Health*. New York: Oxford University Press.

Claridge, G. & Broks, P. (1984). Schizotypy and Hemisphere Function .1. Theoretical Considerations and the Measurement of Schizotypy. *Personality and Individual Differences, 5,* 633-648.

Claridge, G., McCreery, C., Mason, O., Bentall, R., Boyle, G., & Slade, P. (1996). The factor structure of 'schizotypal' traits: A large replication study. *British Journal of Clinical Psychology, 35*, 103-115.

Clarke, S., Bellmann, A., DeRibaupierre, F., & Assal, G. (1996). Non-verbal auditory recognition in normal subjects and brain-damaged patients: Evidence for parallel processing. *Neuropsychologia*, *34,* 587-603.

Clarke, S., Bellmann, A., Meuli, R. A., Assal, G., & Steck, A. J. (2000). Auditory agnosia and auditory spatial deficits following left hemispheric lesions: evidence for distinct processing pathways. *Neuropsychologia, 38,* 797-807.

Coltheart, M. (1981). The Mrc Psycholinguistic Database. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology, 33,* 497-505.

Crow, T. J. (1980). Molecular Pathology of Schizophrenia - More Than One Disease Process. *British Medical Journal, 280,* 66-68.

Crow, T. J. (1995). A Continuum of Psychosis, One Human Gene, and Not Much Else - the Case for Homogeneity. *Schizophrenia Research, 17,* 135-145.

Crowder, R. G. (1989). Imagery for Musical Timbre. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 472-478.

Cui, X., Jeter, C. B., Yang, D. N., Montague, P. R., & Eagleman, D. M. (2007). Vividness of mental imagery: Individual variability can be measured objectively. *Vision Research, 47,* 474-478.

Cummings, A., Ceponiene, R., Koyama, A., Saygin, A. P., Townsend, J., & Dick, F. (2006). Auditory semantic networks for words and natural sounds. *Brain Research, 1115,* 92-107.

Danion, J. M., Kazes, M., Huron, C., & Karchouni, N. (2003). Do patients with schizophrenia consciously recollect emotional events better than neutral events? *American Journal of Psychiatry, 160,* 1879-1881.

Derenzi, E. & Lucchelli, F. (1994). Are Semantic Systems Separately Represented in the Brain - the Case of Living Category Impairment. *Cortex, 30,* 3-25.

DEsposito, M., Detre, J. A., Aguirre, G. K., Stallcup, M., Alsop, D. C., Tippet, L. J. et al. (1997). A functional MRI study of mental image generation. *Neuropsychologia, 35,* 725-730.

Devlin, J. T., Moore, C. J., Mummery, C. J., Gorno-Tempini, M. L., Phillips, J. A., Noppeney, U. et al. (2002). Anatomic constraints on cognitive theories of category specificity. *Neuroimage, 15,* 675-685.

Dick, F., Saygin, A. P., Galati, G., Pitzalis, S., Bentrovato, S., D'Amico, S. et al. (2007). What is involved and what is necessary for complex linguistic and nonlinguistic auditory processing: Evidence from functional magnetic resonance imaging and lesion data. *Journal of Cognitive Neuroscience, 19,* 799-816.

Dierks, T., Linden, D. E. J., Jandl, M., Formisano, E., Goebel, R., Lanfermann, H. et al. (1999). Activation of Heschl's Gyrus during auditory hallucinations. *Neuron, 22,* 615-621.

Doehrmann, O., Naumer, M. J., Volz, S., Kaiser, J., & Altmann, C. F. (2008). Probing category selectivity for environmental sounds in the human auditory brain. *Neuropsychologia, 46,* 2776-2786.

Drake, R., Haddock, G., Tarrier, N., Bentall, R., & Lewis, S. (2007). The Psychotic Symptom Rating Scales (PSYRATS): Their usefulness and properties in first episode psychosis. *Schizophrenia Research, 89,* 119-122.

Eardley, A. F. & Pring, L. (2006). Remembering the past and imagining the future: A role for nonvisual imagery in the everyday cognition of blind and sighted people. *Memory, 14,* 925-936.

Edwards, J., Jackson, H. J., & Pattison, P. E. (2002). Emotion recognition via facial expression and affective prosody in schizophrenia: A methodological review. *Clinical Psychology Review, 22,* 789-832.

Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex, 7,* 181-192.

Engelien, A., Stern, E., & Silbersweig, D. (2001). Functional neuroimaging of human central auditory processing in normal subjects and patients with neurological and neuropsychiatric disorders. *Journal of Clinical and Experimental Neuropsychology, 23,* 94-120.

Escher, S., Romme, M., Buiks, A., Delespaul, P., & van Os, J. (2002). Independent course of childhood auditory hallucinations: A sequential 3-year follow-up study. *Schizophrenia Research, 53,* 4.

Eysenck, H. J. & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire*. London: Hodder and Stoughton.

Farah, M. J. & Smith, A. F. (1983). Perceptual Interference and Facilitation with Auditory Imagery. *Perception & Psychophysics, 33,* 475-478.

Forde, E. M. E., Francis, D., Riddoch, M. J., Rumiati, R. I., & Humphreys, G. W. (1997). On the links between visual knowledge and naming: A single case study of a patient with a category-specific impairment for living things. *Cognitive Neuropsychology, 14,* 403-458.

Fossati, A., Raine, A., Carretta, H., Leonardi, B., & Maffei, C. (2003). The three-factor model of schizotypal personality: invariance across age and gender. *Personality and Individual Differences, 35,* 1007-1019.

Fox, P. T., Miezin, F. M., Allman, J. M., Vanessen, D. C., & Raichle, M. E. (1987). Retinotopic Organization of Human Visual-Cortex Mapped with Positron-Emission Tomography. *Journal of Neuroscience, 7,* 913-922.

Freedman, R., Adler, L. E., Olincy, A., Waldo, M. C., Ross, R. G., Stevens, K. E. et al. (2002). Input dysfunction, schizotypy, and genetic models of schizophrenia. *Schizophrenia Research, 54,* 25-32.

Freeman, D. & Garety, P. A. (2003). Connecting neurosis and psychosis: the direct influence of emotion on delusions and hallucinations. *Behaviour Research and Therapy, 41,* 923-947.

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage, 6,* 218-229.

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *Neuroimage, 19,* 1273-1302.

Frith, C. D. & Done, D. J. (1988). Towards A Neuropsychology of Schizophrenia. *British Journal of Psychiatry, 153,* 437-443.

Funnell, E. & Sheridan, J. (1992). Categories of Knowledge - Unfamiliar Aspects of Living and Nonliving Things. *Cognitive Neuropsychology, 9,* 135-153.

Gaab, N., Gabrieli, J. D. E., & Glover, G. H. (2007a). Assessing the influence of scanner background noise on auditory processing. I. An fMRI study comparing three experimental designs with varying degrees of scanner noise. *Human Brain Mapping, 28,* 703-720.

Gaab, N., Gabrieli, J. D. E., & Glover, G. H. (2007b). Assessing the influence of scanner background noise on auditory processing. II. An fMRI study comparing auditory processing in the absence and presence of recorded scanner noise using a sparse design. *Human Brain Mapping, 28,* 721-732.

Gaffan, D. & Heywood, C. A. (1993). A Spurious Category-Specific Visual Agnosia for Living Things in Normal Human and Nonhuman-Primates. *Journal of Cognitive Neuroscience, 5,* 118-128.

Gainotti, G. (2000). What the locus of brain lesion tells us about the nature of the cognitive defect underlying category-specific disorders: A review. *Cortex, 36,* 539-559.

Gainotti, G. & Silveri, M. C. (1996). Cognitive and anatomical locus of lesion in a patient with a category-specific semantic impairment for living beings. *Cognitive Neuropsychology, 13,* 357-&.

Garety, P. A., Kuipers, E., Fowler, D., Freeman, D., & Bebbington, P. E. (2001). A cognitive model of the positive symptoms of psychosis. *Psychological Medicine, 31,* 189-195.

Golden, R. R. & Meehl, P. E. (1979). Detection of the Schizoid Taxon with Mmpi Indicators. *Journal of Abnormal Psychology, 88,* 217-233.

Goldner, E. M., Hsu, L., Waraich, P., & Somers, J. M. (2002). Prevalence and incidence studies of schizophrenic disorders: A systematic review of the literature. *Canadian Journal of Psychiatry-Revue Canadienne de Psychiatrie, 47,* 833-843.

Gottesman, I. I. & Shields, J. (1982). *Schizophrenia and Genetics: A Twin Study Vantage Point*. London: Academic Press.

Gould, L. N. (1949). Auditory Hallucinations and Subvocal Speech - Objective Study in A Case of Schizophrenia. *Journal of Nervous and Mental Disease, 109*, 418-427.

Grasby, P. M., Frith, C. D., Friston, K. J., Simpson, J., Fletcher, P. C., Frackowiak, R. S. J. et al. (1994). A Graded Task Approach to the Functional Mapping of Brain-Areas Implicated in Auditory Verbal Memory. *Brain, 117,* 1271-1282.

Green, M. F. & Kinsbourne, M. (1990). Subvocal Activity and Auditory Hallucinations - Clues for Behavioral Treatments. *Schizophrenia Bulletin, 16,* 617-625.

Gygi, B., Kidd, G. P. K., & Watson, C. S. (2007). Similarity and categorization of environmental sounds. *Perception & Psychophysics, 69,* 839-855.

Haddock, G., McCarron, J., Tarrier, N., & Faragher, E. B. (1999). Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS). *Psychological Medicine, 29,* 879-889.

Haddock, G., Slade, P. D., Prasaad, R., & Bentall, R. P. (1996). Functioning of the phonological loop in auditory hallucinations. *Personality and Individual Differences, 20*, 753-760.

Hall, D., Goncalves, M., Summerfield, Q., Foster, J., & Bowtell, R. (2000). Quantification of scanner noise effects in auditory fMRI. *British Journal of Audiology, 34,* 95.

Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R. et al. (1999). "Sparse" temporal sampling in auditory fMRI. *Human Brain Mapping, 7,* 213-223.

Halpern, A. R. (1988). Mental Scanning in Auditory Imagery for Songs. *Journal of Experimental Psychology-Learning Memory and Cognition, 14,* 434-443.

Halpern, A. R. & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex, 9,* 697-704.

Halpern, A. R., Zatorre, R. J., Bouffard, M., & Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia, 42,* 1281-1292.

Hart, J. & Gordon, B. (1992). Neural Subsystems for Object Knowledge. *Nature, 359,* 60-64.

Hattiangadi, N., Pillion, J. P., Slomine, B., Christensen, J., Trovato, M. K., & Speedie, L. J. (2005). Characteristics of auditory agnosia in a child with severe traumatic brain injury: A case report. *Brain and Language*, *92*, 12-25.

Hemsley, D. R. (1993). A Simple (Or Simplistic-Questionable) Cognitive Model for Schizophrenia. *Behaviour Research and Therapy, 31,* 633-645.

Hickok, G., Love, T., Swinney, D., Wong, E. C., & Buxton, R. B. (1997). Functional MR imaging during auditory word perception: A single-trial presentation paradigm. *Brain and Language, 58,* 197-201.

Hoshiyama, M., Gunji, A., & Kakigi, R. (2001). Hearing the sound of silence: a magnetoencephalographic study. *Neuroreport, 12,* 1097-1102.

Hubbard, T. L. (2010). Auditory Imagery: Empirical Findings. *Psychological Bulletin, 136,* 302-329.

Hulleman, J. & Humphreys, G. W. (2007). Maximizing the power of comparing single cases against a control sample: An argument, a program for making comparisons, and a worked example from the Pyramids and Palm Trees Test. *Cognitive Neuropsychology, 24,* 279-291.

Humphreys, G. W. & Forde, E. M. E. (2001). Category specificity in mind and brain? *Behavioral and Brain Sciences, 24,* 497-509.

Intons-Peterson, M. J. (1980). The Role of Loudness in Auditory Imagery. *Memory & Cognition, 8,* 385-393.

Intons-Peterson, M. J., Russell, W., & Dressel, S. (1992). The Role of Pitch in Auditory Imagery. *Journal of Experimental Psychology-Human Perception and Performance, 18,* 233-240.

Intonspeterson, M. J. (1980). The Role of Loudness in Auditory Imagery. *Memory & Cognition, 8,* 385-393.

Intonspeterson, M. J., Russell, W., & Dressel, S. (1992). The Role of Pitch in Auditory Imagery. *Journal of Experimental Psychology-Human Perception and Performance, 18*, 233-240.

Jensen (2005). Auditory Imagery: A review and challenges ahead. *Technical Report SSKKII, Center for Cognitive Science, Goteberg, Sweden.*

Johnson, M. K. (1997). Source monitoring and memory distortion. *Philosophical Transactions of the Royal Society B-Biological Sciences, 352*, 1733-1745.

Johnson, M. K. & Raye, C. L. (1981). Reality Monitoring. *Psychological Review, 88*, 67-85.

Johnson, M. K., Taylor, T. H., & Raye, C. L. (1977). Fact and Fantasy - Effects of Internally Generated Events on Apparent Frequency of Externally Generated Events. *Memory & Cognition, 5,* 116-122.

Jolicoeur, P. & Kosslyn, S. M. (1985). Is Time to Scan Visual Images Due to Demand Characteristics. *Memory & Cognition, 13,* 320-332.

Kaas, J. H. & Hackett, T. A. (2008). The Functional Neuroanatomy of the Auditory Cortex. In I.B.Allan, K. Akimichi, M. S. Gordon, W. Gerald, D. A. Thomas, H. M. Richard, D. Peter, O. Donata, F. Stuart, K. B. Gary, M. C. Bushnell, H. K. Jon, & G. Esther (Eds.), *The Senses: A Comprehensive Reference* (pp. 765-780). New York: Academic Press.

Kaney, S. & Bentall, R. P. (1989). Persecutory Delusions and Attributional Style. *British Journal of Medical Psychology, 62*, 191-198.

Kendler, K. S., Mcguire, M., Gruenberg, A. M., & Walsh, D. (1995). Schizotypal Symptoms and Signs in the Roscommon Family Study - Their Factor Structure and Familial Relationship with Psychotic and Affective-Disorders. *Archives of General Psychiatry, 52*, 296-303.

Kendler, K. S., Thacker, L., & Walsh, D. (1996). Self-report measures of schizotypy as indices of familial vulnerability to schizophrenia. *Schizophrenia Bulletin, 22,* 511-520.

Kensinger, E. A. & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition, 31*, 1169-1180.

Kerns, J. G. (2005). Positive schizotypy and emotion processing. *Journal of Abnormal Psychology, 114,* 392-401.

Keshavan, M. S., Montrose, D. M., Pierri, J. N., Dick, E. L., Rosenberg, D., Talagala, L. et al. (1997). Magnetic resonance imaging and spectroscopy in offspring at risk for schizophrenia: Preliminary studies. *Progress in Neuro-Psychopharmacology & Biological Psychiatry, 21*, 1285-1295.

Kety, S. S., Rosentha, D., Wender, P. H., & Schulsin, F. (1971). Mental Illness in Biological and Adoptive Families of Adopted Schizophrenics. *American Journal of Psychiatry, 128,* 302-&.

Kolinsky, R., Fery, P., Messina, D., Peretz, I., Evinck, S., Ventura, P. et al. (2002). The fur of the crocodile and the mooing sheep: A study of a patient with a category-specific impairment for biological things. *Cognitive Neuropsychology, 19,* 301-342.

Kosslyn, S. M. (1973). Scanning Visual Images - Some Structural Implications. *Perception & Psychophysics, 14,* 90-94.

Kosslyn, S. M. (1994). *Image and Brain: The Resolution of the Imagery Debate.* Cambridge, MA: MIT Press.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience, 2,* 635-642.

Krabbendam, L., Myin-Germeys, I., Hanssen, M., de Graaf, R., Vollebergh, W., Bak, M. et al. (2005). Development of depressed mood predicts onset of psychotic disorder in individuals who report hallucinatory experiences. *British Journal of Clinical Psychology, 44*, 113-125.

Kraemer, D., Macrae, C. N., Green, A., & Kelley, W. (2005). The sound of silence: Musical imagery reactivates primary auditory cortex. *Journal of Cognitive Neuroscience,* 235.

Kraepelin, E. (1919). *Dementia Praecox and Paraphrenia*. Edinburgh: E.& S. Livingstone.

Kraut, M. A., Pitcock, J. A., Calhoun, V., Li, J., Freeman, T., & Hart, J. (2006). Neuroanatomic organization of sound memory in humans. *Journal of Cognitive Neuroscience, 18,* 1877-1888.

Lambert, S., Sampaio, E., Scheiber, C., & Mauss, Y. (2002). Neural substrates of animal mental imagery: calcarine sulcus and dorsal pathway involvement - an fMRI study. *Brain Research, 924,* 176-183.

Laroi, F., Marczewski, P., Van der Linden, M., & Evrard, M. (2003). The emotional content of material and reality monitoring function in subjects predisposed towards hallucinations. *Brain and Cognition, 51,* 196-198.

Laroi, F., Van der Linden, M., & Marczewski, P. (2004). The effects of emotional salience, cognitive effort and meta-cognitive beliefs on a reality monitoring task in hallucination-prone subjects. *British Journal of Clinical Psychology, 43,* 221-233.

Launay, G. & Slade, P. (1981). The Measurement of Hallucinatory Predisposition in Male and Female Prisoners. *Personality and Individual Differences, 2,* 221-234.

Lehmann, S. & Murray, M. M. (2005). The role of multisensory memories in unisensory object discrimination. *Cognitive Brain Research, 24,* 326-334.

Levitan, C., Ward, P. B., Catts, S. V., & Hemsley, D. R. (1996). Predisposition toward auditory hallucinations: The utility of the Launay-Slade hallucination scale in psychiatric patients. *Personality and Individual Differences, 21*, 287-289.

Lewis, J. W., Brefczynski, J. A., Phinney, R. E., Janik, J. J., & Deyoe, E. A. (2005). Distinct cortical pathways for processing tool versus animal sounds. *Journal of Neuroscience, 25*, 5148-5158.

Li, C. S. R., Chen, M. C., Yang, Y. Y., Chen, M. C., & Tsay, P. K. (2002). Altered performance of schizophrenia patients in an auditory detection and discrimination task: exploring the 'self-monitoring' model of hallucination. *Schizophrenia Research, 55*, 115-128.

Liddle, P. F. (1987). The Symptoms of Chronic-Schizophrenia - A Reexamination of the Positive-Negative Dichotomy. *British Journal of Psychiatry, 151*, 145-151.

Lissauer, H. (1988). A Case of Visual Agnosia with A Contribution to Theory. *Cognitive Neuropsychology, 5*, 157-192.

MacCallum, R. C., Zhang, S. B., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7*, 19-40.

Maher, B. A. (1974). Delusional Thinking and Perceptual Disorder. *Journal of Individual Psychology, 30*, 98-113.

Marks, D. F. (1973). Visual Imagery Differences in Recall of Pictures. *British Journal of Psychology, 64,* 17-24.

Mason, O. & Claridge, G. (2006). The Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE): Further description and extended norms. *Schizophrenia Research, 82,* 203-211.

Mason, O., Claridge, G., & Jackson, M. (1995). New Scales for the Assessment of Schizotypy. *Personality and Individual Differences, 18*, 7-13.

Maxwell, S. E. & Delaney, H. D. (1993). Bivariate Median Splits and Spurious Statistical Significance. *Psychological Bulletin, 113,* 181-190.

Mcguigan, F. J. (1966). Covert Oral Behavior and Auditory Hallucinations. *Psychophysiology, 3,* 73-&.

McGuire, P. K., Silbersweig, D. A., Murray, R. M., David, A. S., Frackowiak, R. S. J., & Frith, C. D. (1996a). Functional anatomy of inner speech and auditory verbal imagery. *Psychological Medicine, 26,* 29-38.

McGuire, P. K., Silbersweig, D. A., Wright, I., Murray, R. M., Frackowiak, R. S. J., & Frith, C. D. (1996b). The neural correlates of inner speech and auditory verbal imagery in schizophrenia: Relationship to auditory verbal hallucinations. *British Journal of Psychiatry, 169,* 148-159.

Mckelvie, S. J. & Demers, E. G. (1979). Individual-Differences in Reported Visual-Imagery and Memory Performance. *British Journal of Psychology*, *70,* 51-57.

Mechelli, A., Price, C. J., Friston, K. J., & Ishai, A. (2004). Where bottom-up meets top-down: Neuronal interactions during perception and imagery. *Cerebral Cortex*, *14,* 1256-1265.

Meehl, P. E. (1962). Schizotaxia, Schizotypy, Schizophrenia. *American Psychologist*, *17,* 827-838.

Mendez, M. F. (2001). Generalized auditory agnosia with spared music recognition in a left-hander. Analysis of a case with a right temporal stroke. *Cortex*, *37,* 139-150.

Merckelbach, H. & van de Ven, V. (2001). Another White Christmas: fantasy proneness and reports of 'hallucinatory experiences' in undergraduate students. *Journal of Behavior Therapy and Experimental Psychiatry*, *32*, 137-144.

Mintz, S. & Alpert, M. (1972). Imagery Vividness, Reality Testing, and Schizophrenic Hallucinations. *Journal of Abnormal Psychology*, *79*, 310-&.

Morrison, A. P., Frame, L., & Larkin, W. (2003). Relationships between trauma and psychosis: A review and integration. *British Journal of Clinical Psychology*, *42*, 331-353.

Morrison, A. P. & Haddock, G. (1997). Cognitive factors in source monitoring and auditory hallucinations. *Psychological Medicine*, *27,* 669-679.

Morrison, A. P., Wells, A., & Nothard, S. (2002). Cognitive and emotional predictors of predisposition to hallucinations in non-patients. *British Journal of Clinical Psychology, 41,* 259-270.

Morrison, R. L., Bellack, A. S., & Mueser, K. T. (1988). Deficits in Facial-Affect Recognition and Schizophrenia. *Schizophrenia Bulletin, 14,* 67-83.

Mueser, K. T., Bellack, A. S., & Brady, E. U. (1990). Hallucinations in Schizophrenia. *Acta Psychiatrica Scandinavica, 82,* 26-29.

Murray, M. M., Camen, C., Andino, S. L. G., Bovet, P., & Clarke, S. (2006). Rapid brain discrimination of sounds of objects. *Journal of Neuroscience, 26,* 1293-1302.

Myin-Germeys, I., Delespaul, P., & van Os, J. (2005). Behavioural sensitization to daily life stress in psychosis. *Psychological Medicine, 35,* 733-741.

Nagai, M., Kishi, K., & Kato, S. (2007). Insular cortex and neuropsychiatric disorders: A review of recent literature. *European Psychiatry, 22,* 387-394.

Nayani, T. H. & David, A. S. (1996). The auditory hallucination: A phenomenological survey. *Psychological Medicine, 26,* 177-189.

Norman, R. M. G. & Malla, A. K. (1993). Stressful Life Events and Schizophrenia .1. A Review of the Research. *British Journal of Psychiatry, 162,* 161-166.

O'Craven, K. M. & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience, 12,* 1013-1023.

Pastore, R. E., Crawley, E. J., Berens, M. S., & Skelly, M. A. (2003). "Nonparametric" A ' and other modem misconceptions about signal detection theory. *Psychonomic Bulletin & Review, 10,* 556-569.

Peralta, V. & Cuesta, M. J. (2001). How many and which are the psychopathological dimensions in schizophrenia? Issues influencing their ascertainment. *Schizophrenia Research, 49,* 269-285.

Peralta, V., Deleon, J., & Cuesta, M. J. (1992). Are There More Than 2 Syndromes in Schizophrenia - A Critique of the Positive-Negative Dichotomy. *British Journal of Psychiatry, 161,* 335-343.

Peretz, I. & Hyde, K. L. (2003). What is specific to music processing? Insights from congenital amusia. *Trends in Cognitive Sciences, 7,* 362-367.

Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G. et al. (1994). Functional Dissociations Following Bilateral Lesions of Auditory-Cortex. *Brain, 117,* 1283-1301.

Polster, M. R. & Rose, S. B. (1998). Disorders of auditory processing: Evidence for modularity in audition. *Cortex, 34,* 47-65.

Poulton, R., Caspi, A., Moffitt, T. E., Cannon, M., Murray, R., & Harrington, H. (2000). Children's self-reported psychotic symptoms and adult schizophreniform disorder - A 15-year longitudinal study. *Archives of General Psychiatry, 57,* 1053-1058.

Price, C., Thierry, G., & Griffiths, T. (2005). Speech-specific auditory processing: where is it? *Trends in Cognitive Sciences, 9,* 271-276.

Pylyshyn, Z. W. (1973). What Minds Eye Tells Minds Brain - Critique of Mental Imagery. *Psychological Bulletin, 80,* 1-24.

Raine, A. (1991). The Spq - A Scale for the Assessment of Schizotypal Personality Based on Dsm-Iii-R Criteria. *Schizophrenia Bulletin, 17,* 555-564.

Raine, A., Reynolds, C., Lencz, T., Scerbo, A., Triphon, N., & Kim, D. (1994). Cognitive-Perceptual, Interpersonal, and Disorganized Features of Schizotypal Personality. *Schizophrenia Bulletin, 20,* 191-201.

Rankin, P. M. & Ocarroll, P. J. (1995). Reality Discrimination, Reality Monitoring and Disposition Towards Hallucination. *British Journal of Clinical Psychology, 34,* 517-528.

Read, J., van Os, J., Morrison, A. P., & Ross, C. A. (2005). Childhood trauma, psychosis and schizophrenia: a literature review with theoretical and clinical implications. *Acta Psychiatrica Scandinavica, 112,* 330-350.

Reisberg, D., Smith, J. D., Baxter, D. A., & Sonenshine, M. (1989). Enacted Auditory Images Are Ambiguous - Pure Auditory Images Are Not. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology, 41,* 619-641.

Richardson, J. T. E. (1979). Correlations Between Imagery and Memory Across Stimuli and Across Subjects. *Bulletin of the Psychonomic Society, 14,* 368-370.

Roman, R. & Landis, C. (1945). Hallucinations and Mental Imagery. *Journal of Nervous and Mental Disease, 102,* 327-331.

Romme, M. A. J. & Escher, A. D. M. A. (1989). Hearing Voices. *Schizophrenia Bulletin, 15,* 209-216.

Rosentha, D., Wender, P. H., Kety, S. S., Welner, J., & Schulsin, F. (1971). Adopted-Away Offspring of Schizophrenics. *American Journal of Psychiatry, 128,* 307-&.

Sack, A. T., van de Ven, V. G., Etschenberg, S., Schatz, D., & Linden, D. E. J. (2005). Enhanced vividness of mental imagery as a trait marker of schizophrenia? *Schizophrenia Bulletin, 31,* 97-104.

Saha, S., Chant, D., Welham, J., & McGrath, J. (2005). A systematic review of the prevalence of schizophrenia. *Plos Medicine, 2,* 413-433.

Sanjuan, J., Lull, J. J., Aguilar, E. J., Marti-Bonmati, L., Moratal, D., Gonzalez, J. C. et al. (2007). Emotional words induce enhanced brain activity in schizophrenic patients with auditory hallucinations. *Psychiatry Research-Neuroimaging, 154,* 21-29.

Sartoriu, N., Shapiro, R., Barrett, K., & Kimura, M. (1972). Who International Pilot-Study of Schizophrenia. *Psychological Medicine, 2,* 422-&.

Schiffman, J., Ekstrom, M., LaBrie, J., Schulsinger, F., Sorensen, H., & Mednick, S. (2002). Minor physical anomalies and schizophrenia spectrum disorders: A prospective investigation. *American Journal of Psychiatry, 159,* 238-243.

Schneider, K. (1959). *Clinical Psychopathology*. New York: Grune and Stratton.

Schnider, A., Benson, D. F., Alexander, D. N., & Schniderklaus, A. (1994). Nonverbal Environmental Sound Recognition After Unilateral Hemispheric Stroke. *Brain, 117,* 281-287.

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. S. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123*, 2400-2406.

Segal, S. J. & Fusella, V. (1970). Influence of Imaged Pictures and Sounds on Detection of Visual and Auditory Signals. *Journal of Experimental Psychology, 83,* 458-&.

Serper, M. & Berenbaum, H. (2008). The relation between emotional awareness and hallucinations and delusions in acute psychiatric inpatients. *Schizophrenia Research, 101*, 195-200.

Sharma, T., DuBoulay, G., Lewis, S., Sigmundsson, T., Gurling, H., & Murray, R. (1997). The Maudsley family study .1. Structural brain changes on magnetic resonance imaging in familial schizophrenia. *Progress in Neuro-Psychopharmacology & Biological Psychiatry, 21*, 1297-1315.

Sheehan, P. W. (1967). A Shortened Form of Betts Questionnaire Upon Mental Imagery. *Journal of Clinical Psychology, 23,* 386-&.

Shepard, S. & Metzler, D. (1988). Mental Rotation - Effects of Dimensionality of Objects and Type of Task. *Journal of Experimental Psychology-Human Perception and Performance, 14,* 3-11.

Shergill, S. S., Bullmore, E., Simmons, A., Murray, R., & McGuire, P. (2000). Functional anatomy of auditory verbal imagery in schizophrenic patients with auditory hallucinations. *American Journal of Psychiatry, 157,* 1691-1693.

Shergill, S. S., Bullmore, E. T., Brammer, M. J., Williams, S. C. R., Murray, R. M., & McGuire, P. K. (2001). A functional study of auditory verbal imagery. *Psychological Medicine, 31,* 241-253.

Siever, L. J. & Gunderson, J. G. (1983). The Search for A Schizotypal Personality - Historical Origins and Current Status. *Comprehensive Psychiatry, 24,* 199-212.

Sirigu, A., Duhamel, J. R., & Poncet, M. (1991). The Role of Sensorimotor Experience in Object Recognition - A Case of Multimodal Agnosia. *Brain, 114,* 2555-2573.

Skre, I., Onstad, S., Torgersen, S., & Kringlen, E. (1991). High Interrater Reliability for the Structured Clinical Interview for Dsm-Iii-R Axis-I (Scid-I). *Acta Psychiatrica Scandinavica, 84,* 167-173.

Slade, P. & Bentall, R. P. (1988). *Sensory deception: a scientific analysis of hallucination*. Baltimore: John Hopkins University Press.

Smith, B., Fowler, D. G., Freeman, D., Bebbington, P., Bashforth, H., Garety, P. et al. (2006). Emotion and psychosis: Links between depression, self-esteem, negative schematic beliefs and delusions and hallucinations. *Schizophrenia Research, 86,* 181-188.

Smith, J. D. (1992). The auditory hallucinations of schizophrenia. In D.Reisberg (Ed.), *Auditory Imagery* (pp. 151-178). Lawrence Erlbaum Associated, Inc.

Specht, K. & Reul, J. (2003). Functional segregation of the temporal lobes into highly differentiated subsystems for auditory perception: an auditory rapid event-related fMRI-task. *Neuroimage, 20,* 1944-1954.

Spitzer, R. L., Endicott, J., & Gibbon, M. (1979). Crossing the Border Into Borderline Personality and Borderline Schizophrenia - Development of Criteria. *Archives of General Psychiatry, 36,* 17-24.

Stanislaw, H. & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods Instruments & Computers, 31,* 137-149.

Starker, S. & Jolin, A. (1982). Imagery and Hallucination in Schizophrenic-Patients. *Journal of Nervous and Mental Disease, 170,* 448-451.

Stewart, F., Parkin, A. J., & Hunkin, N. M. (1992). Naming Impairments Following Recovery from Herpes-Simplex Encephalitis - Category-Specific. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology, 44,* 261-284.

Strauss, J. S. (1969). Hallucinations and Delusions As Points on Continua Function - Rating Scale Evidence. *Archives of General Psychiatry, 21,* 581-&.

Stuart, G. P. & Jones, D. M. (1996). From auditory image to auditory percept: Facilitation through common processes? *Memory & Cognition, 24,* 296-304.

Tanaka, Y., Nakano, I., & Obayashi, T. (2002). Environmental sound recognition after unilateral subcortical lesions. *Cortex, 38,* 69-76.

Taniwaki, T., Tagawa, K., Sato, F., & Iino, K. (2000). Auditory agnosia restricted to environmental sounds following cortical deafness and generalized auditory agnosia. *Clinical Neurology and Neurosurgery, 102,* 156-162.

Tien, A. Y. (1991). Distributions of Hallucinations in the Population. *Social Psychiatry and Psychiatric Epidemiology, 26,* 287-292.

Tinti, C., Cornoldi, C., & Marschark, M. (1997). Modality-specific auditory imaging and the interactive imagery effect. *European Journal of Cognitive Psychology, 9,* 417-436.

Torrey, E. F. (1987). Prevalence Studies in Schizophrenia. *British Journal of Psychiatry, 150,* 598-608.

van de Ven, V. & Merckelbach, H. (2003). The role of schizotypy, mental imagery, and fantasy proneness in hallucinatory reports of undergraduate students. *Personality and Individual Differences, 35,* 889-896.

Vignolo, L. A. (1982). Auditory Agnosia. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 298,* 49-57.

Vignolo, L. A. (2003). Music agnosia and auditory agnosia - Dissociations in stroke patients. *Neurosciences and Music, 999,* 50-57.

Voisin, J., Bidet-Caulet, A., Bertrand, O., & Fonlupt, P. (2006). Listening in silence activates auditory areas: A functional magnetic resonance imaging study. *Journal of Neuroscience, 26*, 273-278.

Warrington, E. K. & Shallice, T. (1984). Category Specific Semantic Impairments. *Brain, 107,* 829-854.

Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences of the United States of America, 97,* 11125-11129.

Williams, J. B. W., Gibbon, M., First, M. B., Spitzer, R. L., Davies, M., Borus, J. et al. (1992). The Structured Clinical Interview for Dsm-Iii-R (Scid) .2. Multisite Test-Retest Reliability. *Archives of General Psychiatry, 49,* 630-636.

Wu, J. H., Mai, X. Q., Chan, C. C. H., Zheng, Y. Q., & Luo, Y. J. (2006). Event-related potentials during mental imagery of animal sounds. *Psychophysiology, 43,* 592-597.

Yoo, S. S., Lee, C. U., & Choi, B. G. (2001). Human brain mapping of auditory imagery: event-related functional MRI study. *Neuroreport, 12,* 3045-3049.

Young, H. F., Bentall, R. P., Slade, P. D., & Dewey, M. E. (1986). Disposition Towards Hallucination, Gender and Epq Scores - A Brief Report. *Personality and Individual Differences, 7,* 247-249.

Zatorre, R. J. & Halpern, A. R. (1993). Effect of Unilateral Temporal-Lobe Excision on Perception and Imagery of Songs. *Neuropsychologia, 31,* 221-232.

Zatorre, R. J., Halpern, A. R., Perry, D. W., Meyer, E., & Evans, A. C. (1996). Hearing in the mind's ear: A PET investigation of musical imagery and perception. *Journal of Cognitive Neuroscience, 8,* 29-46.

Zubin, J. & Spring, B. (1977). Vulnerability - New View of Schizophrenia. *Journal of Abnormal Psychology, 86,* 103-126.

Appendix A: Instruction to participants for Experiment 1

This experiment aims to investigate vividness of auditory imagery for various types of familiar and unfamiliar sounds, in the following categories: animals, objects and environmental sounds, speech and music.

You will be played a sound and shown the name and picture associated with this sound. Once you have heard the sound, you will be played white noise, then you will see a either the name or picture of the item you just heard.

Please try to imagine the **sound** of the item you just heard (i.e. replay it in your head) and rate how clearly you can imagine this sound. Please use the following rating scale to rate your sound imagery.

Rating Scale
1: No image at all, you only "know" that you are thinking of the object
2: Vague and dim
3: Moderately clear and vivid
4: Clear and reasonably vivid
5: Perfectly clear and as vivid as normal hearing

Take your time to form the image – there is no time limit. If you image is vague, difficult to maintain or confused with other sound images give it a rating of between 1 - 3. If you feel your image is very clear, and you feel you can control the image reliably, give it a rating of between 3 - 5.

You will remain anonymous throughout the study and your results will be kept confidential. You are free to withdraw from this study at any time and your results will be destroyed if you choose to do so.

Feel free to ask any questions you have about the experiment.

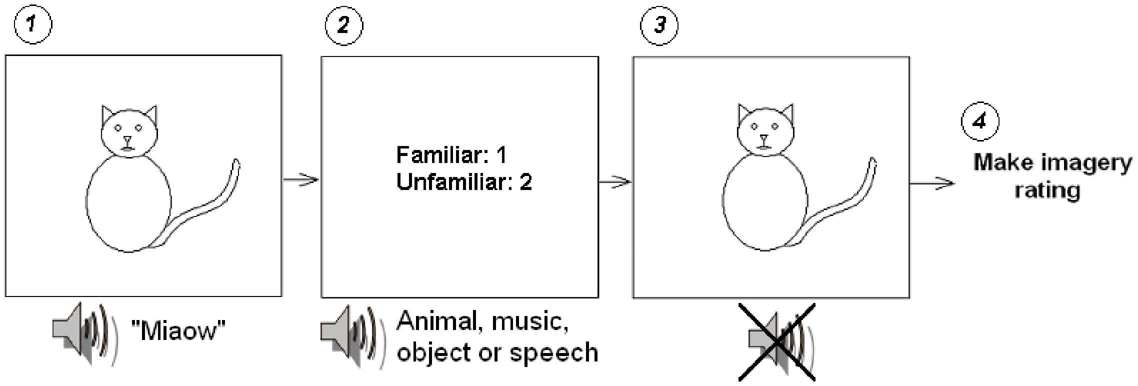
Name:

Signature:

Date:

Appendix B: Instruction to participants for Experiment 2

This experiment aims to investigate imagery for familiar and unfamiliar sounds, in the following categories: **animal sounds, environmental sounds, music and speech**



1. You will be shown a picture, and played the sound that goes with this picture
2. You will then be played another sound which will be familiar or unfamiliar. **If it is familiar to you press '1' and if it is unfamiliar press '2'.**
3. Finally you will be shown the picture shown previously, again.
4. Your task is to try and imagine the sound that goes with the picture, and **rate how clearly you can imagine the sound, using the following scale.**

Rating Scale
1: No image at all, you only "know" that you are thinking of the object
2: Vague and dim
3: Moderately clear and vivid
4: Clear and reasonably vivid
5: Perfectly clear and as vivid as normal hearing

Take your time to form the image – there is no time limit. If you image is vague, difficult to maintain or confused with other sound images give it a rating of between 1 - 3. If you feel your image is very clear, and you feel you can control the image reliably, give it a rating of between 3 - 5.

You will remain anonymous throughout the study and your results will be kept confidential. You are free to withdraw from this study at any time and your results will be destroyed if you choose to do so.

Feel free to ask any questions you have about the experiment.

Please sign below if you understand these terms and are willing to continue with the experiment

Name:

Signature:

Date:

Appendix C: Instruction to participants for Experiment 3a

This experiment aims to investigate sound familiarity and vividness of auditory imagery for various types of familiar and unfamiliar sounds, in the following categories: animals, objects and environmental sounds, speech and music.

In one part of the experiment, you will be played a sound and will be asked to rate how familiar you are with the sound. The sound may be an animal sound, and environmental sound, a tune or a word.

Please rate your familiarity using the following rating scale:

Rating Scale
1: Very unfamiliar
2: Quite unfamiliar
3: Unsure
4: Quite familiar
5: Very familiar

In the second part of the experiment you will be played the same sounds.

This time however, once you have heard the sound, you will be played white noise, and then you will see a rating scale.

When this rating scale appears, please try to imagine the **sound** of the item you just heard (i.e. replay it in your head) and rate how clearly you can imagine this sound. Please use the following rating scale to rate your sound imagery.

Rating Scale
1: No image at all, you only “know” that you are thinking of the object
2: Vague and dim
3: Moderately clear and vivid
4: Clear and reasonably vivid
5: Perfectly clear and as vivid as normal hearing

Take your time to form the image – there is no time limit. If you image is vague, difficult to maintain or confused with other sound images give it a rating of between 1 - 3. If you feel your image is very clear, and you feel you can control the image reliably, give it a rating of between 3 - 5.

You will remain anonymous throughout the study and your results will be kept confidential. You are free to withdraw from this study at any time and your results will be destroyed if you choose to do so.

Feel free to ask any questions you have about the experiment.

Name:

Signature:

Date:

Appendix D: Instruction to participants for Experiment 3b

This experiment will require you to listen to clips of white noise and indicate whether or not you can hear a familiar or unfamiliar **animal sound or music sound** embedded in each one.

In this task there are two conditions: **No Cue and Cue**

No Cue Condition

1. You will first be shown a cross
2. White noise will then be played and a sound may or may not be presented in this white noise
3. Your task is to indicate whether or not you heard the sound in the white noise or not: **press '1' if it was present and '2' if it was absent**.
4. You will also be asked how confident you are that your answer is correct, which you will be asked to rate from 1 to 3: **press keys '1' to '3'.**
5. You will be asked what the sound was. If you did hear a sound, **press 1 if it was an animal, and press 2 if it was music.** If you **did not** hear a sound press 3.
6. You will also be asked whether this sound was familiar or unfamiliar: **Press 1 if it was an familiar, and press 2 if it was unfamiliar.** If you **did not** hear a sound press 3.

Cue Condition

1. You will be played a target sound, and shown a picture associated with this sound.
2. A cross will then appear for 2 seconds.
3. After this you will be shown the picture again and played some white noise. The target sound **may or may not** be played in this white noise.
4. Your task is to indicate whether or not you heard the target sound in the white noise or not: **press '1' if it was present and '2' if it was absent**.
5. You will also be asked how confident you are that your answer is correct, which you will be asked to rate from 1 to 3: **press keys '1' to '3'.**

You will remain anonymous throughout the study and your results will be kept confidential. You are free to withdraw from this study at any time and your results will be destroyed if you choose to do so.

Feel free to ask any questions you have about the experiment.

Name:

Signature:

Date:

Appendix E: Auditory Imagery Questionnaire

This questionnaire investigates ability to form auditory (sound) and visual images.

You will remain anonymous throughout the study and your results will be kept confidential. You are free to withdraw from this study at any time and your results will be destroyed if you choose to do so.

In order to complete this questionnaire please read the rating scale carefully. For each question try to clear your mind and focus on each item to imagine each of the situations described.

Rating Scale
1: No image at all, you only “know” that you are thinking of the object
2: Vague and dim
3: Moderately clear and vivid
4: Clear and reasonably vivid
5: Perfectly clear and as vivid as normal hearing

If you image is vague, difficult to maintain or confused with other sound images give it a rating of between 1 - 3. If you feel your image is very clear, and you feel you can control the image reliably, give it a rating of between 3 - 5.

Imagery for sounds

Imagine the sound of…	**Rating**
a telephone ringing	
a wolf howling	
a police siren	
someone whispering ‘how are you?’	
someone saying ‘hello’	
a piano playing ‘three blind mice’	
a sheep baaing	
the tune of the song ‘happy birthday’	

Imagine you are speaking to a relative or friend you know well and who you frequently speak to and imagine the sound of…

	Rating
The sound of them greeting you	
The tone of their voice	
How they sound when speaking on the phone	
How they sound when pronouncing a French word such as “au revior”	

Imagine you are near a barking dog and imagine the sound of…

	Rating
The loudness of each bark	
The type of bark (i.e. angry, playful, frightened)	
A growl before each bark	
Two dogs barking at each other	

Imagine the hearing the song “Jingle Bells” and imagine the sound of…

	Rating
The tune of the song	
A piano playing the main tune	
Two instruments playing the main tune	
A band playing the main tune	

Imagine someone is vacuuming and imagine the sound of…

	Rating
When first switched on	
The sound of vacuuming	
The sound of a something getting stuck	
The sound of the vacuuming with occasional bumps against furniture	

Appendix F: Instructions to participants for Experiment 7a

This experiment aims to investigate vividness of auditory imagery for positive (e.g. “nice”), negative (e.g. “destroy”) or neutral (e.g. “clue”) spoken words.

You will be played a spoken word and then you will be played white noise. After this you will see a rating scale.

When this rating scale appears, please try to imagine the **sound** of the spoken word you just heard (i.e. replay it in your head) and rate how clearly you can imagine this word being spoken. Please use the following rating scale to rate your sound imagery.

Rating Scale
1: No image at all, you only “know” that you are thinking of the word
2: Vague and dim
3: Moderately clear and vivid
4: Clear and reasonably vivid
5: Perfectly clear and as vivid as normal hearing

Take your time to form the image – there is no time limit. If you image is vague, difficult to maintain or confused with other sound images give it a rating of between 1 - 3. If you feel your image is very clear, and you feel you can control the image reliably, give it a rating of between 3 - 5.

You will remain anonymous throughout the study and your results will be kept confidential. You are free to withdraw from this study at any time and your results will be destroyed if you choose to do so.

Feel free to ask any questions you have about the experiment.

Name:

Signature:

Date:

Appendix G: Instructions to participants for Experiment 7b

This experiment will require you to listen to clips of white noise and indicate whether or not you can hear a word embedded in each one. The types of words that may be presented in the white noise will be positive (e.g. “nice”), negative (e.g. “destroy”) or neutral (e.g. “clue”).

In this task there are two conditions: **No Cue and Cue**

No Cue Condition
7. You will first be shown a cross
8. White noise will then be played and a sound may or may not be presented in this white noise
9. Your task is to indicate whether or not you heard the sound in the white noise or not: **press ‘1’ if it was present and ‘2’ if it was absent**.
10. You will also be asked how confident you are that your answer is correct, which you will be asked to rate from 1 to 3: **press keys ‘1’ to ‘3’.**
11. If you responded that the word was present, you will also be asked what type of word you heard. **Press 1 if it was positive, press 2 if it was negative or press 3 if it was neutral**

Cue Condition
6. You will be played the target word
7. A cross will then appear for 2 seconds, during which you should hold the sound of the word being spoken in memory.
8. After this you will played some white noise. Please imagine the target word, and listen to the white noise. The target sound **may or may not** be played in this white noise.
9. Your task is to indicate whether or not you heard the target sound in the white noise or not: **press ‘1’ if it was present and ‘2’ if it was absent**.
10. You will also be asked how confident you are that your answer is correct, which you will be asked to rate from 1 to 3: **press keys ‘1’ to ‘3’.**

You will remain anonymous throughout the study and your results will be kept confidential. You are free to withdraw from this study at any time and your results will be destroyed if you choose to do so.

Feel free to ask any questions you have about the experiment.

Name:

Signature:

Date:

Appendix H: Further Analyses of acoustic characteristics of the sounds

The following analyses determined how the acoustic characteristics of the target sounds in the sound detection experiments, influenced their detection in noise. Here analyses focused on two acoustic properties: the Root Mean Squared power intensity and Harmonic to Noise Ratio. Previous studies have also calculated these measures to determine the effect of acoustic properties of sounds in neuroimaging tasks (Lewis et al., 2005; Murray, Camen, Andino, Bovet, & Clarke, 2006).

RMS is a measure of power associated with each sound (in dB). the harmonic to noise ratio is a measure of the amount of harmonic acoustic material present in a sound source, compared to non-harmonic noise. This measure quantifies the acoustic energy of the harmonic content of a sound (Lewis et al., 2005). The higher the HNR the less noise there is in the sound, for instance white noise has a HNR of -8 dB whereas a simple tone has a HNR of 85 dB. This factor may therefore influence detection of a sound in noise, because the further away the HNR is from noise, the easier it may be to detect.

Cool Edit Pro was used to calculate RMS for each sound at the threshold level of intensity. The average The PRAAT sound analysis tool (www.fon.hum.uva.nl/pratt) calculated the HNR for each sound at the threshold level

Firstly ANOVAs were conducted to determine the affects of vividness and familiarity on variations of RMS and HNR.

Differences in RMS and HNR dependent on sound vividness and familiarity

Analysis of the RMS for each sound consisted of a two-way ANOVA. The factors were vividness (low vs. high) and familiarity (familiar vs. unfamiliar). This revealed no significant main effect of vividness, $F(1, 124) = 0.900$, $p = 0.345$, *partial* $\eta 2 = 0.007$, but a significant main effect of familiarity, $F(1, 124) = 33.731$, $p < 0.001$, *partial* $\eta 2 = 0.214$, showing a higher RMS for familiar sounds. There was also no significant interaction between vividness and familiarity.

Analysis of HNR consisted of a two-way ANOVA. Again the factors were vividness (low vs. high) and familiarity (familiar vs. unfamiliar). This revealed a no significant main effect of vividness, $F(1, 124) = 0.396$, $p = 0.530$, *partial* $\eta 2 = 0.003$, but a significant main effect of familiarity, $F(1, 124) = 14.980$, $p < 0.001$, *partial* $\eta 2 = 0.108$, showing a higher HNR for familiar sounds. There was also no significant interaction between vividness and familiarity.

A number of analyses were then conducted for each Experiment.

Experiment 3b: Further Analysis

Correlations established the association between the signal detection measures (c, d' and confidence ratings) and the acoustic measures (RMS and HNR).

ANOVA and ANCOVA analyses were then conducted on each sound detection measure, to establish the influence of the HNR on sound detection.

Analysis One: Correlations with acoustic measures

The proportion of hits, false positives and average confidence ratings for hits and false positives were calculated for each sound, as well as the criterion and d' measures for the cue condition. These factors were then correlated with the HNR and RMS to investigate their possible influence on performance (see Table 31).

Firstly there was a significant correlation between HNR and RMS, $r\ (128) = -0.601$, $p<0.001$, showing that a higher HNR was associated with lower RMS. There was also a significant correlation between HNR and vividness, $r(128) = 0.488$, $p < 0.001$, and between HNR and familiarity, $r(128) = 0.310$, $p <0.001$. Comparisons were not made between sound vividness and familiarity and RMS, because the sounds presented in the previous experiment were not presented at the same intensity level as those presented at threshold in the current experiment.

Table 31: Correlations between acoustic measures and signal detection measures

Measure	Condition	HNR	RMS
Hits	Cue	0.378**	-0.003
	No cue	0.364**	0.103
False positives	Cue	0.201*	-0.231*
Criterion	Cue	-0.120	0.166
D'	Cue	0.217*	0.148
Confidence (Hits)	Cue	0.396**	< 0.001
	No cue	0.307**	0.094
Confidence (FP)	Cue	-0.018	-0.098

*** = p<0.05; ** = p<0.001**

Further analysis of hits, false alarms, criterion, d' and confidence ratings over item, determined how much of an affect the acoustic sound characteristics had on sound detection performance.

Hits and false positives: the average proportion of hits were calculated for each item in each condition, and the average number of false positives were calculated for each item in the cue condition only (the same could not be done for the no cue condition as false positives were made to the category, rather than the specific item).

Criterion and d': these measures were calculated for items in the cue condition only, due to the fact that no cue false positive rates could not be calculated

Confidence ratings (hits and false positives): the average confidence rating for each condition, for each item was calculated for the hits. For the confidence ratings in false positives, the cue condition only was analysed due to the reason mentioned above.

Summary

This analysis showed that HNR significant correlated with all measures of sound detection performance. Analyses three consisted of three-way ANCOVA's, to determine whether HNR explained the affects of vividness familiarity and condition. Firstly, ANOVAs were conducted without covariates. In the analyses of the criterion and d', vividness and familiarity were between participants factors. For hits, false alarms and confidence ratings, an additional factor of condition was included as a repeated measures factor. HNR was then included in the analysis as a covariate.

Analysis Two: ANOVA and ANCOVA analyses over item

Hits and False Positives

Analysis of the hit rate consisted of a three-way split plot ANOVA (between participants factors: vividness, familiarity, within-participants factors: condition).

A significant main effect of vividness, $F(1, 124) = 6.884$, $p = 0.010$, partial $\eta 2 = 0.053$, showed a higher hit rate for high vividness sounds. There was also a significant main effect of condition, $F(1, 124) = 115.046$, $p < 0.001$, partial $\eta 2 = 0.481$, revealing a higher hit rate in the cue condition. There was an interaction between vividness and familiarity, $F(1, 124) = 4.002$, $p = 0.048$, partial $\eta 2 = 0.031$. There was no significant difference between familiar and unfamiliar sounds low in vividness, $F(1, 62) = 0.436$, $p = 0.512$, partial $\eta 2 = 0.007$, but a significantly higher hit rate for unfamiliar compared to familiar sounds high in vividness, $F(1, 62) = 7.258$, $p = 0.009$, partial $\eta 2 = 0.105$.

Criterion

Analysis of the criterion revealed no significant main effects or interactions.

d'

Analysis of the d' data over item revealed a near significant interaction between vividness and familiarity, $F(1, 124) = 3.451$, $p = 0.066$, partial $\eta 2 = 0.027$. This showed no difference between familiar and unfamiliar low vividness sounds, $F(1, 62) = 0.390$, $p = 0.535$, partial $\eta 2 = 0.006$, but greater sensitivity to unfamiliar compared to familiar high vividness sounds, $F(1, 62) = 5.542$, $p = 0.022$, partial $\eta 2 = 0.082$.

Confidence Ratings

Confidence in hits

An analysis over items revealed a borderline effect of vividness, $F(1, 121) = 3.795$, $p = 0.054$, partial $\eta 2 = 0.030$, suggesting that participants had greater confidence when detecting high vividness sounds. There was also a significant main effect of condition, *F(1, 121) = 64.613, p <0.001 , partial ŋ2 = 0.348*, with participants having greater confidence in the cue condition. There was also a significant interaction between vividness and condition, *F(1, 121) = 3.874, p = 0.051, partial ŋ2 = 0.031*; confidence was greater for high compared to low vividness sounds in the cue condition, *F(1, 124) =7.333, p = 0.008, partial ŋ2 = 0.056*, but no difference between low and high vividness sounds in the no cue condition, *F(1, 124)=1.405, p = 0.238 , partial ŋ2 = 0.011*.

Confidence in false positives

Confidence in false positives (for the cue condition only) were analysed over items but there were no significant main effects or interactions.

ANCOVA analyses with acoustic measures

The results of the ANCOVA analyse for each sound detection measure can be found in Table 32 and the highlighted effects are those that differ from the original analysis over item, without HNR.

Table 32: ANCOVA results with HNR as covariate

Measure	Effect	F	Sig.	Partial η^2
Hits (N = 123)	Vividness	1.731	0.191	0.014
	Condition	115.046	<0.001	0.481
	Vividness * Familiarity	7.836	0.006	0.060
FP (N = 123)	no significant effects			
Criterion (N = 123)	no significant effects			
d' (N = 123)	HNR	7.300	0.008	0.056
	Vividness * Familiarity	5.173	0.025	0.040
Conf. in hits (N = 120)	Vividness	3.997	0.048	0.032
	Condition	17.528	<0.001	0.126
	Vividness * Condition	2.561	0.112	0.021
	Vividness * Familiarity	6.400	0.012	0.051
Conf. in FP (N = 104)	no significant effects			

The inclusion of HNR removed the significant effects of vividness on hit rate and the interaction between vividness and condition on the confidence in hits. It also made reliable the interaction between vividness and familiarity on the confidence in false positives, however. For familiar sounds, confidence was greater for low vividness sounds, than high vividness sounds, $F(1, 60) = 7.851, p=0.007$, *partial* $\eta^2 = 0.116$, but there was no difference between high and low vividness sounds for unfamiliar sounds, $F(1, 59) = 0.094, p=0.760$, *partial* $\eta^2 = 0.002$.

Further analyses: Experiment 6b

The following analyses followed the same pattern as that in Experiment 3b. The aim here was to firstly establish the influence of the acoustic characteristics on detection of this larger sample of participants. Also the aim was to determine how these acoustic characteristics varied as a function of hallucination proneness.

Analysis 1: Correlations with acoustic measures

Each of the sound detection factors was correlated with the Harmonic to Noise Ratio (HNR) and the Root Mean Square (RMS) measures, to assess the impact of variance in the acoustic properties of the signals. Full descriptions of the HNR and RMS indices can be found in Chapter 2. The correlations for the high and low hallucination prone groups can be found in Table 33.

Table 33: Correlations between sound detection task measures and vividness, familiarity, HNR and RMS (r-values)

	Measure	Condition	HNR	RMS
	Hits (N = 128)	Cue	0.384**	-0.015
		No Cue	0.392**	0.06
	FP (N = 128)	Cue	0.268**	-0.234*
	Criterion (N = 128)	Cue	-0.371**	0.062
	d' (N = 128)	Cue	0.145	0.168
	Confidence (hits) (N = 128)	Cue	0.360**	-0.011
		No Cue	0.401**	0.031
HHP	**Confidence (FP)** (N = 119)	Cue	-0.016	0.1
	Hits (N = 128)	Cue	0.402**	-0.017
		No Cue	0.324**	0.141
	FP (N = 128)	Cue	0.240*	-0.273*
	Criterion (N = 128)	Cue	-0.421**	0.064
	d' (N = 128)	Cue	0.252*	0.105
	Confidence (hits) (N = 128)	Cue	0.318**	-0.031
		No Cue	0.155	0.09
LHP	**Confidence (FP)** (N = 116)	Cue	0.214*	-0.208*

HHP = high hallucination prone; LHP = low hallucination prone
*** = p<0.05, ** = p<0.001**

HNR: For high hallucination prone participants, increased HNR of the sounds was associated with increased hit rate, false positive rate and confidence in hits, and decreased criterion. It was not related to d’ or confidence in false positives. For low hallucination prone participants, increased HNR was associated with increased hit rate, false positive rate, d’ and confidence in hits, and decreased criterion, but was not related to false positives.

RMS: for the high hallucination prone group, higher RMS was associated with decreased false positive rate only, but for low hallucination prone group, higher RMS was associated with decreased false positive rate, and decreased confidence in false positive responses.

Analysis Two: ANOVA analyses over item

Criterion

Analysis of the criterion was calculated for items in the cue condition only, due to the fact that false positive rates could not be calculated for the no cue condition. A three way split plot (within participants factor: hallucination proneness group, between participants factors: vividness and familiarity) was conducted on the data over items. Condition could not be analysed here because false positives in the no cue condition were made to a vividness category rather than to any particular item.

A main effect of vividness, $F(1, 124)=12.468, p=0.001$, *partial* $\eta^2=0.091$, showed a significant bias to high vividness sounds. There was also a significant interaction between vividness and familiarity, $F(1, 124)=5.024, p=0.027$, *partial* $\eta^2=0.039$. This showed that there was no difference between low and high vividness sounds when familiar, $t(62)=0.825, p=0.412$, whereas the bias was greater for high vividness sounds, when unfamiliar, $t(62)=4.623, p<0.001$.

There was also a significant interaction between hallucination proneness and familiarity, $F(1, 124)=5.354, p=0.022$, *partial* $\eta^2=0.041$. When the sound was familiar, high hallucination prone participants were more biased, $F(1, 63)=6.578$,

$p=0.013$, partial $\eta^2=0.095$, but when unfamiliar the hallucination proneness groups did not differ, *$F(1, 63)=0.491$, $p=0.486$, partial $\eta^2=0.008$,*

No other main effect or interaction was significant.

d'

Analysis of the d' was calculated for items in the cue condition only, due to the fact that no cue false positive rates could not be calculated. The data were analysed over items using a three-way split-plot analysis (within participants factor: hallucination proneness; between participants factors: vividness and familiarity). A significant interaction between vividness and familiarity, *$F(1, 124)=6.580$, $p=0.012$, partial $\eta^2=0.050$,* was found. When low vividness sounds, there was no difference in sensitivity between familiar and unfamiliar sounds, *$F(1, 62) = 1.739$, $p = 0.192$, partial $\eta^2 = 0.027$*, whereas when the sound high vividness, sensitivity was greater for unfamiliar sounds compared to familiar sounds, *$F(1, 62) = 7.350$, $p = 0.009$, partial $\eta^2 = 0.106$.* No other main effects or interactions were significant.

Confidence Ratings

The average confidence rating for each condition, for each item was calculated for the hits. For the confidence ratings in false positives, the cue condition only was analysed due to the reason mentioned above. A four way split plot ANOVA on the confidence in hits (repeated measures: hallucination proneness and condition, between participants: vividness and familiarity) revealed a significant main effect of vividness, *$F(1, 123) = 3.916$, $p = 0.050$, partial $\eta^2 = 0.031$*. Participants were more confident in having heard high vividness sounds. There was also a significant main effect of condition, *F(1,123) = 73.092, p <0.001 , partial* η^2 *= 0.373*, showing that participants were more confident that they heard a sound in the cue condition. There was a three-way interaction between vividness, familiarity and condition, *F(1, 123) = 4.184, p = 0.043, partial* η^2 *= 0.033*. In the cue condition, the interaction between vividness and familiarity was not reliable, *F(1, 123) = 1.881, p = 0.173, partial* η^2 *= 0.015*. In the no cue condition, the interaction between vividness and familiarity was significant, *F(1, 123) = 7.800, p = 0.006, partial* η^2 *= 0.060*. Further analysis showed no significant difference between familiar and unfamiliar low vividness sounds, *F(1, 61) = 1.958, p = 0.167, partial* η^2 *= 0.031*, but for high vividness sounds, participants were significantly more confident in hearing unfamiliar sounds. There was also a significant interaction between hallucination proneness, vividness and condition, *F(1, 123) = 4.883, p = 0.029, partial* η^2 *= 0.038.* In the cue condition, the interaction between vividness and hallucination proneness was not significant, *(1, 126) = 0.364, p = 0.548, partial* η^2 *= 0.003.* In the no cue condition however the vividness - hallucination proneness interaction was reliable, *F (1, 125) = 9.840, p = 0.002, partial* η^2 *= 0.073*. When the sound was low in vividness, high and low hallucination prone participants did not differ in their confidence for detecting sounds, *F (1, 62) = 3.455, p = 0.068, partial* η^2 *= 0.053*, but when the sound was high in vividness, the high hallucination prone group were more confident in the sounds presence, *F (1, 63) = 8.725, p = 0.004, partial* η^2 *= 0.122.*

The confidence ratings on false positive responses for the cue condition were analysed using, a three way split plot ANOVA (within participants factor: hallucination proneness; between participants factor: vividness and familiarity). A significant main effect of hallucination proneness was found, *F(1,105) = 4.072, p = 0.046, partial* η^2 =

0.037. High hallucination prone participants were more confident in their false positive responses. There was also a significant three-way interaction between hallucination proneness, vividness and familiarity, $F(1, 105) = 6.299$, $p = 0.014$, partial $\eta^2 = 0.057$. For familiar sounds, the interaction between vividness and hallucination proneness was significant, $F(1, 50) = 12.565$, $p = 0.001$, partial $\eta^2 = 0.201$. For low vividness sounds, high hallucination prone participants were more confident that they heard a sound than low hallucination prone participants, $F(1, 22) = 7.860$, $p = 0.010$, partial $\eta^2 = 0.263$. There was no significant interaction between vividness and hallucination proneness for unfamiliar sounds, $F(1, 55) = 0.060$, $p = 0.807$, partial $\eta^2 = 0.001$. No other main effects or interactions were significant.

Hits (cue and no cue) and false positives (cue condition only)

As the criterion and d' measures could not be calculated over items for the no cue condition, we also analysed the hit rate for each item in each condition, so as to gain an idea of sensitivity to the target over item, in both conditions.

A four way split plot ANOVA (within-participants factors: hallucination proneness group and condition; between participants factors: vividness and familiarity) was conducted on the hit data over item.

The main effect of vividness was significant, $F(1, 124) = 7.090$, $p = 0.009$, partial $\eta^2 = 0.054$, showing that more hits were made to high vividness sounds. There was also a significant main effect of hallucination proneness, $F(1, 124) = 23.503$, $p < 0.001$, partial $\eta^2 = 0.159$. High hallucination prone participants heard more sounds than low hallucination prone participant. The main effect of condition, $F(1, 124) = 123.877$, *p*

<0.001 , partial η^2 *= 0.500,* showed that significantly more sounds were heard in the cue condition compared to the no cue condition. The interaction between vividness and condition, $F(1, 124) = 6.376$, $p = 0.013$, *partial* $\eta^2 = 0.049$, was also significant. In the cue condition, significantly more hits were made to high vividness sounds, $F(1, 126) = 10.704$, $p = 0.001$, *partial* $\eta^2 = 0.078$, but in the no cue condition, there was no significant difference between the low and high vividness sounds, $F(1, 126) = 3.340$, $p = 0.070$, *partial* $\eta^2 = 0.026$. There was also a significant interaction between vividness and familiarity, $F(1, 124) = 3.976$, $p = 0.048$, *partial* $\eta^2 = 0.031$. For low vividness sounds, there was no significant difference between familiar and unfamiliar sounds, $F(1, 62) = 0.467$, $p = 0.497$, *partial* $\eta^2 = 0.007$, but for high vividness sounds, significantly more hits were made to unfamiliar sound. Finally there was an interaction between hallucination proneness and familiarity, $F(1, 124) = 10.950$, $p = 0.001$, *partial* $\eta^2 = 0.081$. For familiar sounds, high hallucination prone participants made more hits than low hallucination prone participants, $F(1, 63) = 32.558$, $p < 0.001$, *partial* $\eta^2 = 0.341$, however for unfamiliar sounds, there was no significant difference between the hallucination proneness groups, $F(1, 63) = 1.241$, $p = 0.269$, *partial* $\eta^2 = 0.019$.

For the false positive responses, a three way split plot ANOVA was conducted (within-participants factor: hallucination proneness; between participants factors: vividness and familiarity). There was a significant main effect of hallucination proneness, $F(1, 124) = 5.868$, $p = 0.017$, *partial* $\eta^2 = 0.045$. High hallucination prone participants made more false positive responses than low hallucination prone individuals. The main effect of vividness was also significant, $F(1, 124) = 8.856$, $p = 0.004$, *partial* $\eta^2 = 0.067$. More false positives were made to high vividness sounds.

ANCOVA analyses with acoustic measures

The majority of the sound detection measures correlated significantly with the HNR (see correlational analysis below). Therefore HNR was used as a covariate in the same ANOVA analyses as reported above, for all signal detection measures. Table 34 shows the significant results of these ANCOVA's. The highlighted results are those that have changed from significant to not significant or vice versa.

Table 34: ANCOVA results with HNR as covariate

Measure	Effect	F	Sig.	Partial η^2
	HNR	23.130	<0.001	0.158
	Vividness	2.016	0.158	0.016
	Condition	48.867	<0.001	0.284
	Hal. Prone.	7.065	0.009	0.054
	Vividness * Condition	9.341	0.003	0.071
Hits	Vividness * Familiarity	8.044	0.005	0.061
(N = 123)	Hal. Prone. * Familiarity	10.973	0.001	0.082
	HNR	5.383	0.022	0.042
FP	Vividness	0.208	0.649	0.003
(N = 123)	Hal. Prone.	0.037	0.848	<0.001
	HNR	5.109	0.026	0.040
	Vividness	0.702	0.404	0.006
	Familiarity	4.165	0.043	0.033
Criterion	Vividness * Familiarity	9.794	0.002	0.074
(N = 123)	Hal. Prone * Familiarity	6.163	0.014	0.048
	HNR	9.603	0.002	0.072
	Hal. Prone.	7.296	0.008	0.056
	Hal. Prone. * HNR	5.378	0.022	0.042
d'	Hal. Prone. * Familiarity	4.332	0.039	0.034
(N = 123)	Vividness * Familiarity	9.438	0.003	0.071
	HNR	22.498	<0.001	0.156
	Condition	23.172	<0.001	0.160
	Vividness * Familiarity	9.271	0.003	0.071
	Vividness * Familiarity * Condition	4.543	0.035	0.036
	Hal. Prone. * Condition	8.990	0.003	0.069
Conf. in hits	Hal. prone. * Condition * Vividness	0.042	0.839	<0.001
(N = 122)	Hal. Prone. * Condition * HNR	6.266	0.014	0.049
Conf. in FP	Hal. Prone.	1.295	0.258	0.012
(N = 104)	Hal. prone. * Vividness * Familiarity	5.973	0.016	0.054

Inclusion of HNR as a covariate removed the following effects from the analyses:

- Effect of vividness on hits, false positives and criterion
- Effect of hallucination proneness on false positive rates and confidence in false positives

- The interaction between hallucination proneness, condition and vividness in the confidence in hits

The ANCOVA also made significant the effect of hallucination proneness, and the interaction between hallucination proneness and familiarity on the d' data. It also made significant the effect of familiarity on the criterion. Individuals with high hallucination proneness scores were more sensitive to sounds overall, compared to low hallucination prone participants. High hallucination prone participants showed a trend to being more sensitive to familiar sounds than low hallucination prone participants, $F(1, 62) = 3.084, p = 0.084$, *partial* $\eta^2 = 0.047$, but the two groups did not differ in sensitivity when the sounds were unfamiliar, $F(1, 62) = 1.834, p = 0.178$, *partial* $\eta^2 = 0.014$.

Appendix I: Auditory recognition deficits of a patient with frontal lobe lesions

Auditory recognition disorders have been reported to occur following damage to the temporal lobe and sub-cortical regions. However assessment of such disorders varies, and little research has been conducted into recognition abilities of patients with damage to other brain regions. The current case study is of DS, a patient with frontal lobe lesions who displays auditory recognition deficits. Animal sounds, environmental sounds, music and speech were investigated using several tasks, designed to investigate associative and apperceptive recognition abilities. In addition visual recognition abilities of DS were also investigated using, tests analogous to those used in the auditory domain. DS's performance on each test was compared to 9 age-matched control participants. Pure tone audiometry revealed that DS's hearing was at near normal levels, however he was impaired relative to controls at naming, matching two-exemplars of sounds, sound-picture matching and answering questions about different sounds, for all sound categories. DS had most problems with animal sounds and musical instruments. In addition he was impaired relative to controls at visual analogies of these tasks. DS therefore appears to have a combination of associative and apperceptive impairments in both auditory and visual recognition, which are worse for music instruments and animals.

Auditory recognition deficits of a patient with frontal lobe lesions

Introduction

Auditory agnosia

Neuropsychological research has investigated patients with damage to auditory regions, such as the superior and medial temporal lobes to assess the existence of auditory disorders. This has however been problematic for a number of reasons. Firstly auditory processing disorders often go unnoticed because speech production and processing areas are also damaged, masking other sound impairments (Polster et al., 1998). Also due to the position and structure of the auditory processing areas, damage is relatively rare, and often bilateral temporal lobe damage is necessary for the disorder to be evident (Polster et al., 1998; Engelien, Stern, & Silbersweig, 2001).

Despite the rarity of auditory disorders, impairments analogous to visual processing disorders have been discovered. A number of studies have found dissociations within auditory perception, such as amusia (disordered music perception), pure word deafness (disordered speech perception) and auditory agnosia (disorder sound perception) (Vignolo, 2003; Polster et al., 1998; Hattiangadi et al., 2005; Taniwaki, Tagawa, Sato, & Iino, 2000; Peretz et al., 1994).

Auditory agnosia is an inability to recognise sounds, despite normal or near normal hearing ability (Clarke, Bellmann, Meuli, Assal, & Steck, 2000), and often co-occurs with pure word deafness, though isolated non-verbal auditory agnosia has been documented (Taniwaki et al., 2000). In addition and like visual agnosia, apperceptive (e.g. perceptual knowledge) and associative (e.g. semantic knowledge) impairments have been discovered (Vignolo, 1982).

Regardless of the similarities between auditory and visual agnosia, few studies have looked for inter-modality similarities. Extensive testing of visual agnosia symptoms and dissociations has resulted in a number of established assessments of apperceptive and associative deficit (i.e. VOSP, Warrington and James, 1986; BORB, Riddoch and Humphreys, 1993) that are lacking in auditory agnosia research. Some auditory equivalents of the tests of visual agnosia have been developed (Clarke et al., 2000; Vignolo, 1982; Mendez, 2001), but they have not been systematically investigated or compared to performance in other modalities (Polster et al., 1998).

Category-specific deficits.

Within the general class of visual agnosia it has been well established that patients can have selective deficits for particular types of objects, with perhaps the most common distinction being between impairments for living and non-living items (Caramazza & Shelton, 1998; Warrington et al., 1984). A smaller body of research has investigated category specific deficits across modalities. In Kolinsky et al. (2002) study, category-specific deficits for visual, verbal and auditory input were investigated in a patient with bilateral hippocampal and temporal lobe lesions. The patient was impaired at naming living items (i.e. animals) from vision and audition and in describing items from written words, when compared to non-living items (i.e. objects), suggesting that category-specific deficits may extend across modalities.

Other studies have also found that auditory category-specific impairments are associated with similar disorders in visual recognition (Caramazza et al., 1998; Gainotti et al., 1996). However such investigations tend to focus on patients with known visual deficits, to see whether they also have a deficit with auditory information, rather than vice versa (i.e. do patients with auditory recognition disorder have similar disorders for visual information).

While studies of visual recognition point to a general distinction between living and non-living things, finer-grained dissociations have also been reported (i.e. for tools, food, body parts, and musical instruments; see (Gainotti et al., 1996). This suggests that the living – non-living dissociation is not cut-and-dried, and that there may be finer-grained representation of particular classes of stimulus (see Caramazza and Shelton, 1998). One pertinent distinction in visual recognition disorders is that between musical instruments and other non-living objects, suggesting that they form a separate semantic category (Gainotti et al., 1996; Warrington et al., 1984). Similarly studies of auditory amusia have also suggested that musical impairments can co-occur with auditory agnosia (Hattiangadi et al., 2005), but can also exist as an isolated deficit (Peretz & Hyde, 2003).

The current case study investigates the auditory processing abilities of a patient with aphasia resulting from left frontal lobe damage. The aim was to investigate recognition abilities for different types of non verbal sounds, to determine the possibility of category-specific sound perception deficits. Animal sounds, environmental sound and music stimuli were used in this experiment, in a range of recognition, matching and classification tasks. In addition performance on such tasks was compared to similar tasks in the visual modality, to determine whether the deficits found in sound perception were specific to sounds or existed across modalities. Speech recognition was assessed using a standardised assessment (PALPA).

Method

Patient

DS, a 74-year-old former train inspector, suffered a stroke in 1995. He continued to live at home and functioned in a relatively self-sufficient manner in everyday life. An MRI scan revealed damage to the left inferior, superior, and middle frontal gyri (see Figure 33; also Humphreys and Forde, 1998). He was hemiplegic in his right arm. DS's full IQ score on the WAIS was 72, his General Memory Score on the WMS was 50, and his forward digit span (4) was outside the control range. DS's low-level visual perception and object naming were relatively normal: he scored 50/50 (100%) on unusual views matching using tests from BORB set (Riddoch and Humphreys, 1993), and he named 25/29 (86%) of a common set of everyday objects. His errors were to call a bowl a "jug or cup," a knife a "fork," a lunchbox a "dolly," and a teapot a "kettle." Despite the presence of these few semantic errors in naming, DS used the objects appropriately. However, consistent with his frontal lobe damage, DS performed poorly on the Stroop test, scoring more than 3 SDs away from the control mean on a clinical version of the task (see Humphreys and Forde, 1998).

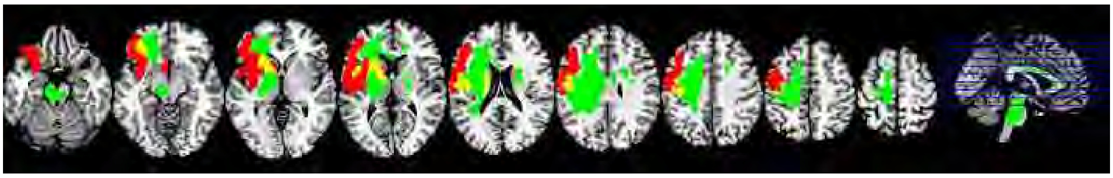


Figure 33: DS's lesions: Left inferior frontal gyrus: pars opercularis, pars triangularis and pars orbitalis, left rolandic operculum, left insula, left middle frontal gyrus, left precentral gyrus, left postcentral gyrus, left caudate and putamen; (red =grey matter; green = white matter; yellow = either grey or white matter).

DS's hearing was assessed using pure tone audiometry. This revealed that he had normal age-related hearing loss at all frequencies (see ISO 7029 for normal hearing levels of different age groups). The audiogram can be found in

Figure 34.

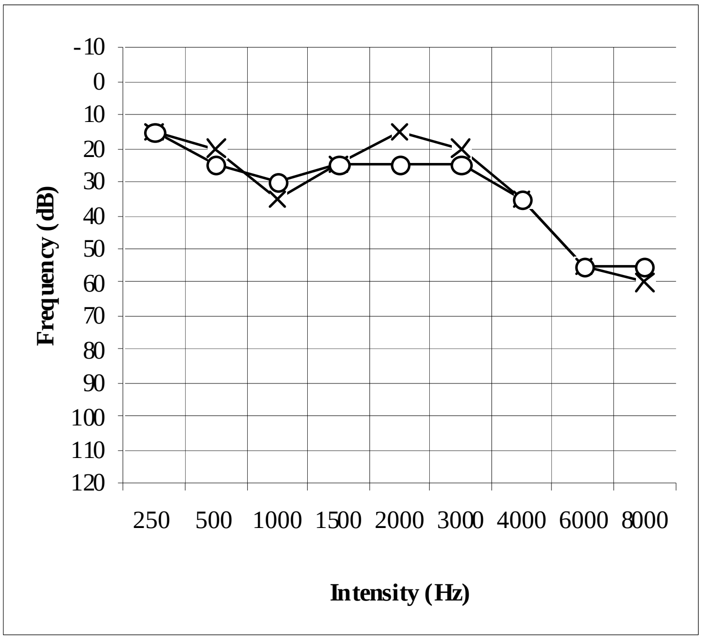


Figure 34: Pure tone threshold audiometry for the left ear (x) and the right ear (o)

Controls

9 male controls participated in this experiment, with a mean age of 64 years (± 7.48). All stated that they did not have any problems with their hearing.

All of the tasks described here were conducted twice with DS and once with each of the control participants. Sound stimuli were presented over standard computer speakers, and were presented using E-prime (Psychology Software Tools, Pittsburgh, PA) (for matching and recognition tasks) and the sound recorder function (for classification and naming tasks).

Stimuli

Animal and environmental sounds: Stimuli from these sound categories were collected from a variety of online sources, and sounds considered most archetypal were selected for inclusion in the following tasks. They were then edited using CoolEditPro to be 2sec. in length, with a sampling rate of 22500 Hz and a bit rate of 16 bits. Due to the fact that people are familiar with a relatively small selection of animal and environmental sounds, many of the sounds were used across different tasks (see Appendices for list of sounds, and tasks they were used in).

17 animal sounds and 22 environmental sounds were rated for familiarity by 11 young controls. Each sound was played once, and participants were required to rate how familiar they were with the sound, from 1 (very unfamiliar) to 5 (very familiar). The average familiarity rating of was 4.005 for animal sounds, and 4.079 for object sounds. Paired-samples t-test revealed no significant difference between animal sounds and environmental sounds, $t(10)=0.501, p=0.627$.

Music instruments: Clips of different musical instruments playing unfamiliar musical pieces were also collected from online sources. They were edited using CoolEditPro to be 3.5 sec. in length, with the same sampling and bit rate as for animal and environmental sounds. Due to the fact that many instruments sound quite similar (i.e. French horn and a tuba) a small set of distinct instruments were selected and used across different tasks (see Appendices for list of sounds, and tasks they were used in). The musical instrument sounds were also rated for familiarity in the same way at the non-musical sounds, by the same eleven participants. The average familiarity rating was 4.137, and these ratings did not differ from either animal sounds, $t(10) = 0.756$, $=0.467$, or object sounds, $t(10) = 0.302$, $=0.769$.

Experiment 2a. Animal and object sounds

Method

Naming Sounds

Participants were asked to name the sounds that were used in the following experimental tasks. Due to the fact that only a small number of sounds are easily recognisable, the sounds overlapped between tasks (see Table 36 and Table 37 for a list of sounds, the tasks they were included in and the number of control participants who named them). 17 animal sounds and 22 object sounds were named. Sounds were randomly presented using the "Sound Recorder" function in Windows.

DS named each sound once and between three and nine control participants named each sound. The proportion of control participants who correctly named each sound was calculated and then averaged across the sounds.

Naming Pictures

Participants were asked to name the pictures that were used in the following experimental tasks and overlapped between tasks. 17 animal and 23 object picture.

Pictures were presented in a random order using PowerPoint, and responses were written down. DS and nine control participants named each picture once.

Matching Tasks

Sounds

Exact matching

Participants were told that they would be presented with two sounds, which would either be exactly the same or completely different. Each sound was presented for 2.5sec, with a 2sec fixation period in between. This was then followed by a screen which said “same or different?” to which participants were required to indicate whether the sounds were the same or different, by pressing one of two keys on the computer keyboard. When the pictures were different, the second picture was randomly selected from the other items from the category (i.e. another animal if the first picture was an animal). This meant that each sound was presented the same number of times as either a first or second sound. 10 animal sounds and 10 object sounds were each presented twice in this task – once where the two sounds were the same and once where they were different. Therefore this task included 40 trials.

Two exemplar matching

Participants were presented with two sounds, to which they were required to make a same/different judgement. However for the “same” trials, rather than the sounds being exactly the same, the second sound was a different version of the first. For example if the first sound were a cat meowing, the second sound would be a different cat meowing, but the correct judgement would be that the two sounds are the same as they are of the same animal. For the “different” trials the second sound was a completely different sound, randomly selected from the other sounds in the experiment.

The same 10 animal sounds and 10 object sounds from the previous experiment were presented in this task. Each was presented twice, once where the two sounds were of the same sound, and once where the two sounds were different. In addition, for the trials were the two sounds were of the same thing, alternate versions of the animal and environmental sound target were also presented.

Pictures

Exact matching

The procedure of this task was the same as that for the sound version. Participants were presented with a picture for 2.5sec and then following a fixation period of 2sec, a second picture was presented for 2.5sec. Participants then had to indicate whether the two pictures were exactly the same or of completely different items. When the pictures were different, the second picture was randomly selected from the other items from the category (i.e. another animal if the first picture was an animal). The pictures were full colour photographs of same items that had been used in the sound version of this task, therefore 10 animal pictures and 10 object pictures (approximately 8cm x 8cm) were each presented twice (once were the second picture was the same, and once were it was different). This ensured that each picture was presented equally as a first or a second picture. D. S completed this task once.

Two exemplar matching

The task was the same as above, but this time participants had to judge whether two pictures were of the same or different animals or objects. The two pictures were different, but could be of the same animal or object (i.e. 'same' judgement) or of two different animals or object (i.e. 'different' judgement). Again the pictures were of the same items that had been used in the sound version of this task. DS and four control participants completed this task once.

Classification tasks

Sounds

In this task, participants were played 12 animal sounds and 12 object sounds using the 'Sound Recorder' function in Windows in a random order, and asked a question about the sound. The questions concentrated on the semantic aspects of the sounds, such as where an animal might be found, or whether and object can be used inside, rather than the acoustic or visual aspects associated with the sounds. Each sound was presented twice: once where the correct answer was "yes" (i.e. for a cat meowing: "Can this animal be kept as a pet?") and once where it was "no" (i.e. for a cat meowing: "Would you expect to find this animal in a zoo"). For each sound category (i.e. animal or object sounds) four questions were presented.

Pictures

This task was similar to the previous task. However participants were presented with picture versions of the sounds used. Therefore 12 animal pictures and 12 object pictures were presented in a random order, using PowerPoint. For each picture participants were asked a question about a semantic aspect of the pictured item and these questions were the same as those that were asked about the sounds, i.e. for a cat meowing: “Can this animal be kept as a pet?”

Recognition Task

This task was adapted and updated from Vignolo (1982) Sound Recognition Task. In this task, participants were presented with a sound, and then four pictures, with their task being to match the sound to the numbered picture associated with it, by pressing the appropriate number on the keyboard. The four pictures represented the target sound (i.e. cat meowing), an acoustically related sound (i.e. a baby crying), a semantically related sound (i.e. a dog barking) or an odd sound, unrelated to the target (i.e. a wasp). This task consisted of 9 animal sound targets and 10 object sound targets; therefore there were a total of 19 trials. The sound was played first (approximately 2 sec in length), and then following a 1 sec fixation period, the four pictures were presented until a response was made.

Results

The percentage of correct responses of DS in each of the tasks and in each session can be found in Table 38. DS’s overall correct responses are analysed and compared to control here, however the consistency between the two sessions is also assessed.

Naming

Figure 35 shows the average percentage of correctly named sounds and pictures.

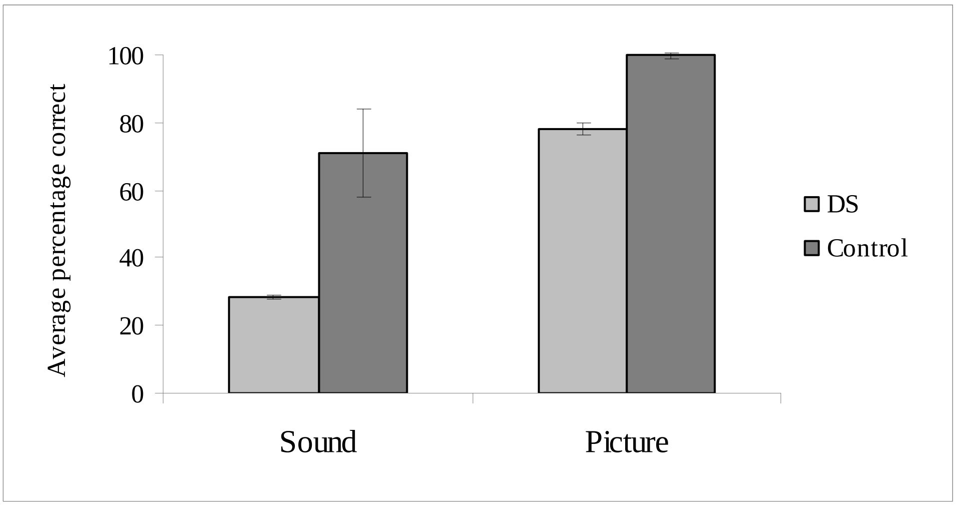


Figure 35: Percentage of correctly named sounds and pictures

Sounds

Summed across the two sessions he performed, DS scored 24/77 (31.12%) correct. He named 9/34 (26.47%) animal sounds correctly and 13/43 (30.23%) object sounds correctly. Performance did not differ between the two sessions (overall 13/38 for session 1 and 11/39 for session 2; *p >0.999)* and the contingency coefficient confirmed that DS performance was consistent across sessions *(c = 0.525, p < 0.001).*

Control participants correctly named 70.83% ± 12.98 of all sounds correctly. Of these they named 69.49% ± 16.15 of animal sounds correctly and 72.86 % ± 11.02 of object sounds correctly.

Pictures

Summed across the two sessions, DS correctly named 61/78 (78.21%) of the pictures correct. He named 20/34 (58.82%) animal pictures correctly and 41/44 (93.18%) object pictures correctly. Performance did not differ between the two sessions (overall 30/39 for session 1 and 31/39 for session 2; McNemar: *p >0.999)* and was also consistent across the sessions, *c = 0.309, p=0.043*.

The control participants named 99.72% ± 1.76 of the pictures correctly. Of these 99.34% ± 2.69 of the animal pictures were named correctly and 100% ± 0 of the object sounds were named correctly.

A four-way loglinear analysis was conducted to compare DS's naming performance to control participants. The factors were participant (DS vs. control), stimuli type (sound vs. picture), stimuli category (animal vs. object) and accuracy (correct vs. incorrect).

A significant interaction between patient and task was found, $\chi^2(1) = 4.293$, $p = 0.038$, showing that DS was significantly worse at naming sounds compared with naming pictures, compared with control participants.

The interaction between task and accuracy revealed that overall, participants were more accurate at naming pictures than sounds, $\chi^2(1) = 24.092$, $p < 0.001$.

There was also a significant interaction between participant and accuracy, $\chi^2(1) = 34.082$, $p < 0.001$, revealing that DS was significantly more inaccurate than control participants.

Matching tasks

Sounds

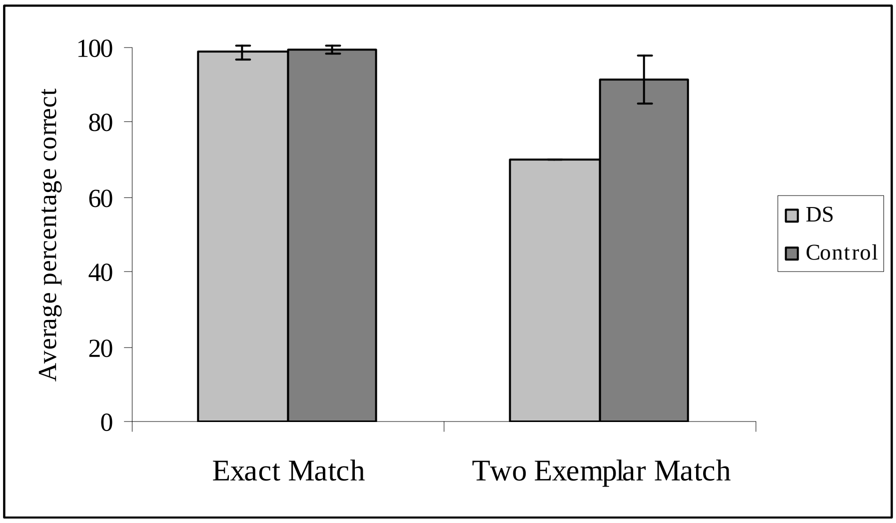


Figure 36: Percentage of correct responses in sound matching tasks

The average percentage of correct responses for DS and each control participants in the two matching tasks can be found in Figure 36.

Exact matching: DS's correctly matched 79/80 (98.75%) across the two sessions. DS correctly matched 39/40 animal sounds (98.75%) and 40/40 object sounds (100%). DS performed at ceiling in one of the sessions (39/40 for session 1 and 40/40 for session 2) therefore tests of consistency could not be performed.

The average performance of the control participants was 39.8/40, (99.5% ± 0.44) and these participants correctly matched 19.89/20 animal sound (99.5% ± 0.33) and 19.89/20 object sounds (99.5% ± 0.33).

DS performed within 3 SD's of the controls performance, however to confirm that performance did not differ relative to the controls, a modified t-test was performed (Hulleman & Humphreys, 2007). This revealed that there was no difference between DS and the control participants overall $F(1, 8) = 0.370, p = 0.299$, nor when animal and object sounds were analysed separately (animal sound: $F(1, 8) = 1.24, p = 0.233$; object sound: $F(1, 8) = 0.10, p = 0.175$).

Two exemplar matching: DS correctly matched 56/80 (70%) of the sounds, across the two sessions. He correctly matched 27/40 (67.5%) animal sounds and 29/40 (72.5%) object sounds. There were 18 errors to 'same' items and 6 errors to 'different' trials. DS's performance did not differ between the two sessions (overall 28/40 for session 1 and 28/40 for session 2; McNemar: $p > 0.999$). In this case the contingency coefficient was not significant suggesting that performance was not consistent between the two sessions, $c = 0.071, p=0.651$.

Average performance of the control participants was 36.5/40 (91.25% ± 2.65), with 18.44/20 animal sounds being matched correctly (92.2% ± 1.51) and 18.11/20 (90.5% ± 1.62) object sounds matched correctly. DS's performance fell below 3 SD's of the control participant's performance. Modified t-tests were performed on the total scores, and the animal and object sound scores, to compare DS and control performance. This revealed that DS was significantly worse than controls overall, $F(1, 8) = 9.39, p <0.001$. Relative to controls he was significantly impaired both at matching animal sounds $F(1, 8) = 9.63, p <0.001$ and object sounds $F(1, 8) = 4.49, p <0.001$, when they were analysed separately.

Pictures

Exact matching: DS correctly matched 68/80 (85%) of the pictures correctly, across the two sessions. He correctly matched 30/40 (75%) of the animal sounds and 38/40 (95%) of the object sounds. Performance did not differ between the two session (overall 32/40 for session 1 and 36/40 for session 2; McNemar: $p = 0.549$). The contingency coefficient assessing the consistency of performance, were not significant *(p = 0.549; c=<0.001, p>0.999)*. Control participants performed at ceiling level.

Two exemplar matching: DS correctly matched 57/80 (71.25%) pictures correctly in across the two session. He correctly matched 25/40 (62.5%) animals correctly and 32/40 (80%) object pictures correctly. Of the errors, 19 were made to ‘same’ items and 6 to ‘different’ items. Performance did not differ between the two sessions (overall 29/40 for session 1 and 28/40 for session 2; McNemar: $p > 0.999$), but it was consistent across the sessions, $c=0.387, p=0.008$.

Control participants correctly matched an average of 38.3/40 (95.8% ± 0.50) sounds. Of these they correctly matched 18.67/20 (93% ± 0.50) animal sounds and 19.67 (98% ± 0.58) of the object sounds.

A three-way loglinear analysis was then performed on DS’s data to assess differences in performance in the matching tasks based on stimuli type (sound vs. picture), category (animal vs. object) and accuracy (correct vs. incorrect). For the exact matching task, there was a significant interaction between stimuli type and accuracy, $\chi^2(1) = 5.897, p = 0.015$, showing that DS was significantly more accurate in the sound matching than the picture matching version. For the two exemplar matching task, only the main effect of accuracy was significant, $\chi^2(1) = 13.165, p <0.001$, showing that DS got more of the comparisons correct than incorrect.

Classification tasks

Figure 37 shows the average performance of DS's and control participants in the sound and picture versions of the classification task. This shows that DS performed similarly whether the item was a sound or a picture, and that DS performance was lower than that of the control participants

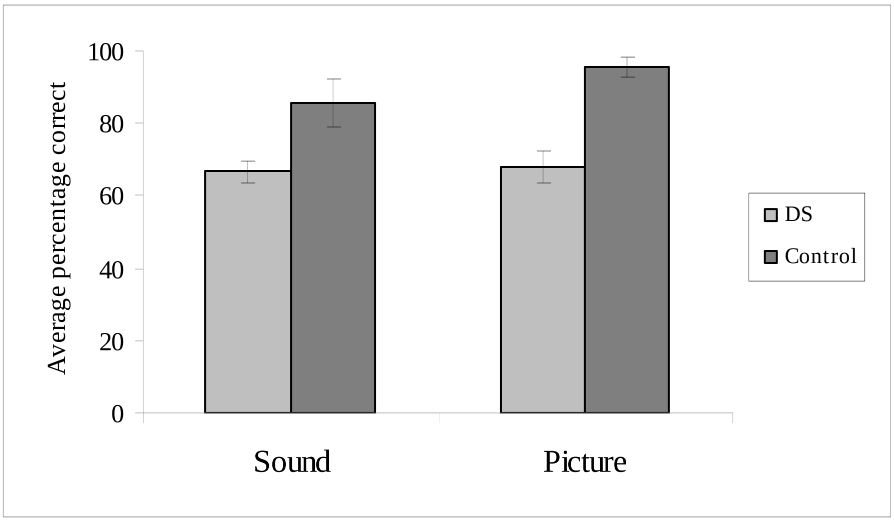


Figure 37: Percentage of correct responses on sound and picture versions of the classification task

Sounds

DS correctly answered 64/96 (66.67%) of the questions across the two sessions. DS correctly answered 28/48 (58.33%) questions about animal sounds and 36/48 (75%) questions about object sounds. Performance between the two sessions did not differ (overall 33/48 for session 1 and 31/48 for session 2; McNemar: $p = 0.38$), however the contingency was also not significant, $c = 0.105$, $p = 0.464$, suggesting his performance was not consistent between the two sessions.

Control participants correctly answered an average of 41.1/48 (85.6% ± 3.10) questions correctly. 20.44/24 (85% ± 2.40) questions about animal sounds were answered correctly and 20.67/24 (86.13% ± 1.73) questions about object sounds were answered correctly.

Again DS performed outside of 3 SD's of the control participant's performance. Modified t-test confirmed that DS was significantly impaired at this task overall, compared to the control participants $F(1, 8) = 7.77, p <0.001$. He was impaired both when animal and object sounds were analysed separately (animal: $F(1, 8) = 6.46, p <0.001$; object: $F(1, 8) = 2.14, p = 0.024$).

Pictures

DS correctly answered 65/96 (67.71%) questions across the two sessions. DS correctly answered 26/48 (54.2%) questions about animals and 39/48 (81.25%) questions about objects. DS's performance did not differ across the sessions (34/48 for session 1, 31/48 for session 2; McNemar: $p < 0.581$), and was consistent across items, $c = 0.361, p=0.007$.

Control participants correctly answered an average of 46/48 (95.8% ± 1.32) questions correctly, with 23.4/24 (95.8% ± 0.88) correctly answered about animals and 22.56/24 (93.96% ± 0.88) about objects.

DS performed outside of 3 SD's of the control participant's performance. Modified t-tests revealed that the difference between DS's overall performance and the averaged control participants was significant $F(1, 8) = 93.73, p <0.001$. Further, it was revealed that the DS and controls differed significantly for both animal sounds ($F(1, 8) = 12.76, p <0.001$) and object sounds ($F(1, 8) = 10.80, p <0.001$).

DS's performance on the sound and picture versions of this task was assessed using a three-way loglinear analysis, with the factors being stimuli type (sound vs. picture), stimuli category (animals vs. objects), and accuracy (correct vs. incorrect). This revealed a significant main effect of accuracy only, $\chi 2\ (1) = 14.338,\ p = 0.026$, showing that DS answered more questions correctly than incorrectly. No other main effects or interactions were significant.

Recognition Task

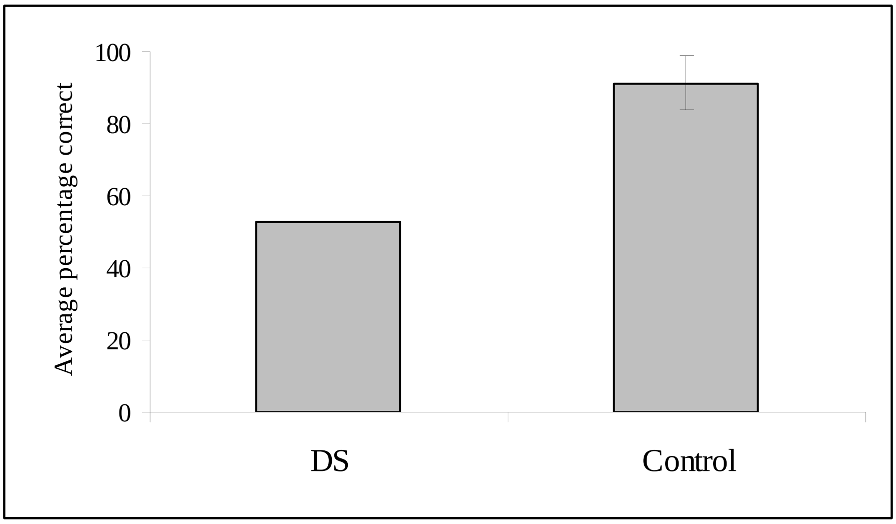


Figure 38: Percentage of correct responses on sound recognition task

DS correctly matched 20/38 (52.63%) pictures to the sounds, across the two sessions. He matched 10/18 (55.55%) animal sounds to the correct picture, and 10/20 (50%) object sounds to the correct picture. Of the errors made, half were acoustic and half were semantic errors. DS's performance did not differ between the two sessions (10/19 for session 1, 10/19 for session 2; McNemar: $p > 0.999$). In this case his performance was not consistent across items *(c = 0.344, p=0.11).*

Control participants by contrast, matched an average of 17.33/19 (91.2% ± 1.4) sounds correctly, with 8.67/9 (96.3% ± 0.71) animals matched to the correct picture and 8.56/10 (85.6% ± 0.88) objects to the correct picture.

DS performed more than 3 SD's below the average of the control participants (see Figure 38). A modified t-test confirmed that DS was significantly impaired compared to controls at the task overall $F(1, 8) = 24.18, p < 0.001$. He was also significantly impaired when animal sounds and object sounds were analysed separately (animal: $F(1, 8) = 24.24, p < 0.001$; object: $F(1, 8) = 14.62, p < 0.001$).

Loglinear analysis was conducted on DS's data, with the factors being stimuli category (animal vs. object) and accuracy (correct vs. incorrect). This revealed no significant main effects or interactions.

DS was also asked to do this task when the sounds were presented at the same time as the pictures. On the first session DS correctly matched 13/14 (92.86%) of the sounds correctly. He made one semantic error to an environmental sound (responding to a picture of a car, when the sound was a helicopter). However due to a technical error, five responses were not recorded. A second assessment using this task however, revealed that DS matched only 9/19 (47.36%) of the sounds correctly.

Comparison between tests

DS's performance on all of the tests was compared. A four-way loglinear analysis was conducted and the factors in this analysis were task (naming, matching, classification, recognition), stimuli type (sound vs. picture), stimuli category (animal vs. object), , and accuracy (correct vs. incorrect).

A significant interaction between task, stimuli and accuracy was found, $\chi^2(1) = 22.54$, $p<0.001$. This interaction reflects the fact that DS's performance was worse on naming than on the classification and matching tasks, particularly for sounds.

A significant interaction between stimuli, category and accuracy was also found, $\chi^2(1) = 4.989$, $p = 0.026$. This interaction arose because DS's performance was worse on animals compared to objects, with this effect being stronger for pictures.

Discussion

The tasks employed here aimed to investigate different aspects of DS's auditory recognition, and how these abilities are related to his visual recognition. Firstly DS was poorer at naming sounds relative to controls, whereas DS did not differ from controls in his ability to name pictures. Both DS and controls were poorer at naming sounds than pictures. DS's naming performance did not differ between animals and objects however.

The matching tasks were performed to assess the existence of apperceptive or associative impairments which may explain DS naming impairments. The exact matching task (i.e. assessment of apperceptive problems) revealed that DS was unimpaired at the sound version of the task, but showed some impairment on the picture version of the task, relative to controls, particularly to animal stimuli. The two exemplar task (i.e. assessment of associative problems) showed that DS was impaired relative to controls on both the sound and picture versions of the matching task.

The classification task also assessed associative abilities by investigating how well DS could relate sounds and pictures to semantic knowledge about animal and object sounds. DS was again impaired at both sound and picture versions of the task relative to controls, and was particularly poor at answering questions about animals.

The recognition task aimed to assess DS's ability to associate sounds with pictures. DS was impaired at this task, relative to controls, but his performance did not differ between animal and object sounds. In addition there was no difference in the number of acoustic and semantic errors. This therefore suggests that though DS has problems matching sounds to their associated pictures, this deficit is not specifically due to either apperceptive or associative impairments.

Finally DS's performance on the naming tasks, the two matching tasks and the classification tasks for animal and environmental sounds and pictures, were compared. DS's performance with pictures was similar between all tasks, but for sounds, he was less accurate at the naming tasks than the other three tasks. In addition for the picture tasks, DS was more inaccurate for animals than object pictures, whereas for sounds there was no difference between animal and object sounds.

These tasks therefore suggest that DS has more of an associative impairment for animals and objects, which is particularly apparent for animal pictures compared to object pictures or sounds. It is also interesting to note that, at least for naming, DS tended to be consistent (there was some inconsistency across items, but this arose on 2 AFC tasks, and inconsistency could result from guessing (a 50% probability of being right or wrong by chance). Item-specific consistency in naming fits with an argument that there is a central semantic deficit.

The following experiment investigates similar tasks, but with musical instruments and speech.

Experiment 2b: Musical instruments and Speech

Naming

Sounds

Table 39 shows the music instrument sound clips that were used in the following tasks. Eighteen 3.5sec musical instrument sound clips were presented in a random order, and played using the “Sound Recorder” function in Windows. Each was named by DS and nine control participants once.

Pictures

Picture versions of the 18 music instruments were also used in the following tasks. These pictures measured approximately 8cm x 8cms and each was named once by DS and nine controls. The pictures were presented in a random order using PowerPoint.

Matching Tasks

Sounds

Exact matching

The procedure of this task was the same as for the animal and object sound task, i.e. participants had to indicate whether the two sound clips were exactly the same or completely different. In this task however, participants were presented with pairs of musical clips (each clip for 3.5secs) with a 2sec fixation period between the two clips. The participants tasks was to judge whether the sounds were the same (i.e. same instrument and same tune) or different (i.e. different instrument and tune).

Eighteen musical instrument clips were presented twice: once where the second clip was exactly the same, and once were it was completely different. Therefore this task included a total of 36 trials.

Two exemplar matching

In this task, participants were again presented with two music clips, to which they had to indicate whether the sounds were the same or different. However for the “same” trials, participants were presented with two different tunes that were played on the same instrument. Therefore the task was to judge whether the tunes were played on the same instrument or not, regardless of what the actual tune was. Eleven musical instrument clips were presented twice: once where the second clip was a different tune, but the same instrument, and once were it was both a different tune and instrument. Therefore this task included a total of 22 trials.

Pictures

Exact matching

This task was the same as that for the sounds; however participants had to make same/different judgements to pictures rather than sounds.

Two exemplar matching

This task was the same as that for the sounds; however participants had to make same/different judgements to pictures rather than sounds.

Tune Recognition

In this task, participants were played two tunes played on a piano, and their task was to indicate whether the tunes were exactly the same, or different. On half the trials the two tunes were the same. In the remaining trials the two sounds were different, however the difficulty level of these trials differed: a third of these trials were at an easy level, with three notes differing between the two tunes, a third at an intermediate level, with two notes differing between the two tunes and the remaining third of the trials being at a difficult level, differing by only one note.

Classification tasks

Sounds

In this task 12 music clips were presented and one of four questions was asked about the sound, focusing on how the instrument was played (i.e. “do you blow this instrument?”). Again each music clip was played twice, once with a positive answer and once with a negative answer

Pictures

In this task 12 pictures of musical instruments were presented twice (once with a positive answer and once with a negative answer), and one of four questions was asked about the sound, focusing on how the instrument was played (i.e. do you blow this instrument?)

Recognition Task

In this task, participants were presented with a tune for 3.5sec, and then 2sec of fixation, followed by two pictures of musical instruments, with their task being to indicate which musical instrument the tune had been played on. This task consisted of 18 trials.

Speech

Speech recognition abilities were assessed using subscales of the Psycholinguistic Assessments of Language Processing in Aphasia (PALPA). The two tests were chosen as those that most closely reflected the tasks used for the other sound types, namely the spoken word-picture matching task and a spoken word matching task.

Spoken word-picture matching task

This task was therefore similar to the sound recognition tasks used for the other sound categories. The patient listened to a word spoken by the experimenter (with the mouth covered to avoid lip reading), and then was asked to match it to one of three line drawn pictures. One of these pictures was the target, a second was semantically related, and a third was acoustically related.

Spoken word matching task

This task was similar to the sound matching tasks used for the other sound categories. This task required DS to determine whether two spoken words were the same or different to each other. Again the words were spoken by the experimenter (with the mouth was covered). When different, the two words could differ in the placement of the difference and the frequency of the word.

Results

The percentage of correct responses of DS in each of the tasks and in each session can be found in

Table **40**. Like the assessment of animals and objects, DS's average correct responses are analysed and compared to control here, and the consistency between the two sessions is also assessed.

Naming

Figure 39 shows the average percentage of correctly named musical instrument sounds and picture. This shows that DS was impaired relative to controls, at naming both sounds and pictures, and that controls were marginally better at naming pictures compared to sounds.

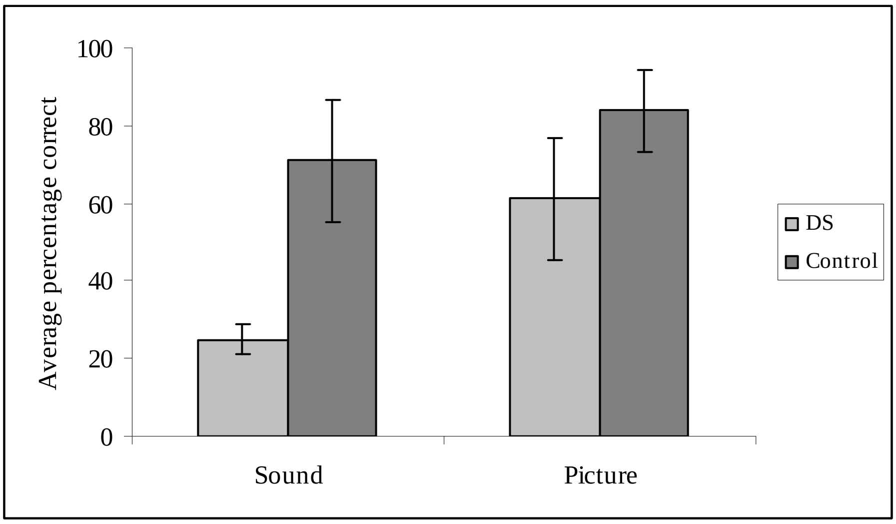


Figure 39: Percentage of correctly named musical instrument sounds and pictures

Sounds

DS correctly identified 9/36 (25%) instruments correctly, across the two sessions. He named 3/36 (8.33%) musical instruments correctly, but mimed 6/36 (16.67%) of the instruments. Performance did not differ across the two sessions, (5/18 for session 1, 4/18 for session 2; McNemar: $p > 0.999$), and the contingency coefficient suggests that the performance was consistent across these session *(c = 0.491, 0.017)*.

Control participants correctly named an average of 70.99% (± 15.66) of the musical instruments correctly.

Pictures

DS correctly identified 21/36 (58.33%) of the musical instruments across the two sessions. He named 5/18 (13.89%) and mimed 16/18 (88.89%) musical instruments correctly. Performance did not differ between the two sessions (session 1 9/18, session 2 12/18; McNemar: $p = 0.51$), however the contingency coefficient was also not significant, suggesting performance was not consistent between sessions, $c = <0.001, p>0.999$.

Control participants correctly named 83.95% (± 10.56) of the musical instruments correctly.

A three-way loglinear analysis was conducted to compare DS performance to controls. The factors were participant type (DS and control), stimuli type (sound vs. picture) and accuracy (correct vs. incorrect). A significant interaction between task and accuracy, $\chi^2(1) = 4.580$, $p = 0.032$, revealed that participants were more accurate at naming pictures than sounds. An interaction between participant and accuracy, $\chi^2(1) = 9.135$, $p = 0.003$, revealed that control participants were more accurate than DS.

Naming performance of musical instruments was then compared to that of naming animals and objects. The score for animal and objects was summed and a loglinear analysis was conducted, with the factors being participant (DS vs. control), stimuli category (musical instruments vs. animals and objects), stimuli type (sounds vs. pictures) and accuracy (correct vs. incorrect).

This revealed a significant interaction between stimuli type and patient, $\chi^2(1) = 4.481$, $p = 0.034$, showing that DS was significantly poorer at naming sounds than control participants, but did not differ to controls in picture naming.

Participants was also significantly less accurate at naming sounds than pictures, $\chi^2(1) = 36.125$, $p < 0.001$.

DS was also less accurate than control participants, $\chi^2(1) = 33.013$, $p < 0.001$.

Matching Tasks

Sounds

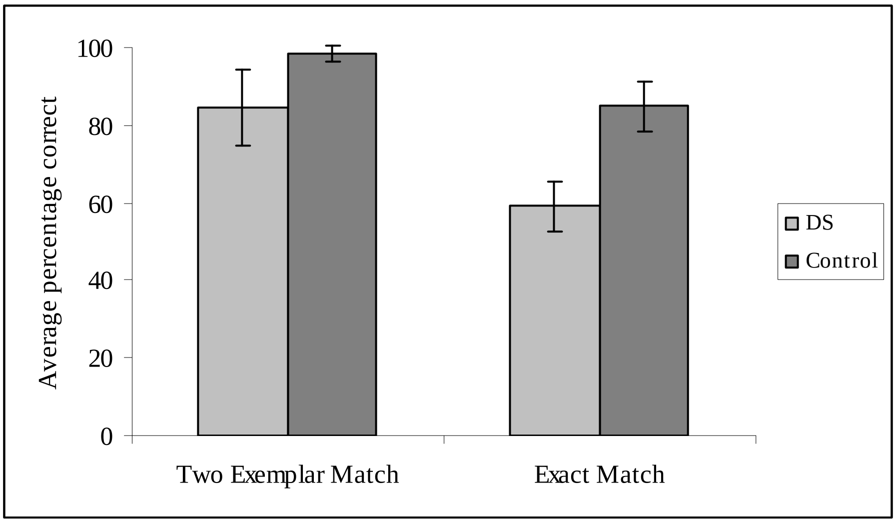


Figure 40: Percentage of correct responses in instrument matching tasks

Exact matching: DS correctly matched 61/72 (84.72%) music clips across the two sessions, and performance did not differ between the two sessions (session 1 28/36, session 2 33/36; McNemar: $p = 0.23$), and was consistent across these sessions, $c = 0.627, p<0.001$.

Control participants matched 35.44/40 (98% ± 0.73). DS's performance was below 3 SD's of the control participants (see Figure 40). Modified t-test revealed that DS performed significantly lower than the control participants $F(1, 8) = 41.44, p = 0.042$.

Two exemplar matching: DS correctly matched 26/44 (59%) of the music clips over the two sessions, with more errors being made to 'same' pairs, than 'different' pairs. Performance did not differ across the two sessions (14/22 for session 1, 12/22 for session 2; McNemar: $p = 0.73$), however it was also not consistent across these sessions $(c=0.251, p=0.225)$.

In contrast, control participants correctly matched 18.67/22 (84% ± 1.41). DS performed outside of 3 SD's of the control participants' performance. Modified t-tests confirmed that this difference was significant $F(1, 8) = 14.74, p < 0.001$.

Pictures

Exact matching: DS correctly matched 68/72 (94.4%) pictures correctly across the two sessions. Consistency between sessions could not be assessed as DS performed at ceiling in one of the sessions (32/36 for session 1, 36/36 for session 2). Control participants also performed at ceiling in this task.

Two exemplar matching: DS correct matched 28/44 (63.6%) pictures correctly across the two sessions. Performance did not differ between the two sessions (15/22 for session 1, 13/22 for session 2; McNemar: $p = 0.73$), but was also not consistent across these sessions, $c = 0.22, p = 0.29$.

In contrast five controls were tested on this task, and they correctly matched 20.2/22 (91.8% ± 4.02). A modified t-test revealed that the performance of DS was significantly poorer, $F(1, 4) = 1.98, p = 0.021$, than controls.

Loglinear analysis was then performed on DS's data to assess differences in performance in the matching tasks based on stimuli type. For the exact matching task, there was a borderline interaction between task and accuracy, $\chi^2(1) = 3.748, p = 0.053$, showing that DS was significantly more accurate in the picture matching than the sound matching version.

For the two exemplar matching task, there were no significant main effects or interactions.

Tune Recognition

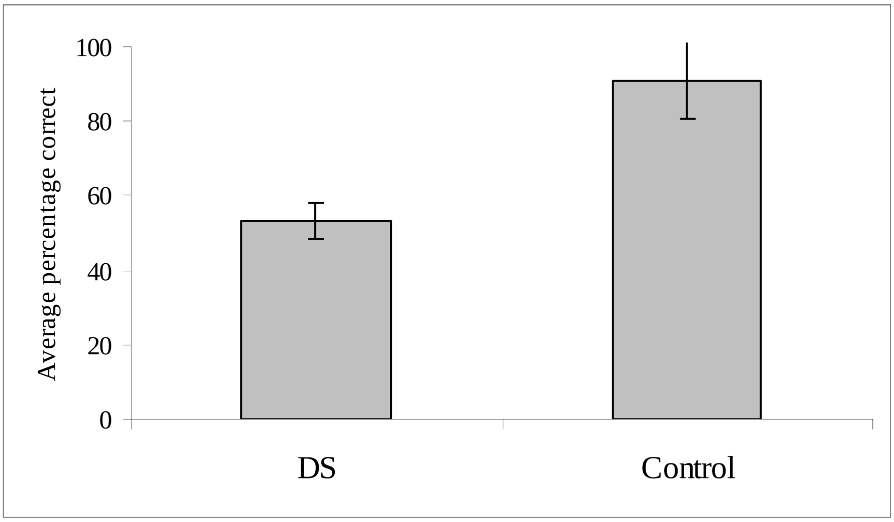


Figure 41: Percentage of correct responses on tune matching task

DS correctly recognised 32/60 (53.33%) tunes across the two sessions, with the majority of errors due to responding ‘different’ on ‘same’ trials. Performance did not differ between the two sessions (15/30 for session 1, 17/30 for session 2; McNemar: $p = 0.75$), and there was a trend towards consistency between the two sessions, $c = 0.319$, $p = 0.065$.

In contrast, control participants correctly matched 27.3/30 (91% ± 3.08) tunes. DS again performed below 3 SD of the mean of the control participants’ performance (see Figure 41). Modified t-tests confirmed this difference between DS and control participants $F(1, 8) = 12.17$, $p < 0.001$.

Classification tasks

Sounds

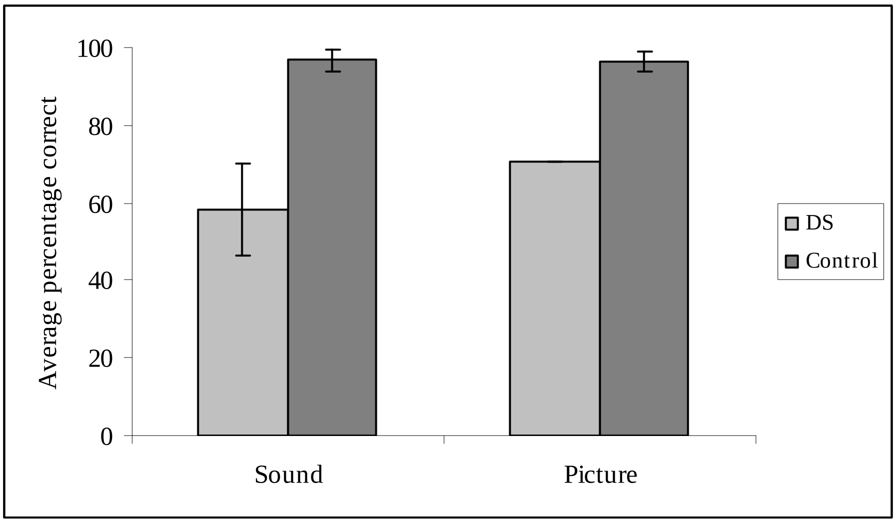


Figure 42: Percentage of correct responses on sound and picture versions of the instrument classification task

DS correctly answered 28/48 (58.33%) of the questions across the two sessions, whereas control participants correctly answered an average of 23.22/24 (96.8% ± 0.67) questions correctly. DS's performance did not differ between the two sessions (12/24 for session 1, 16/24 for session 2; McNemar: $p = 0.42$), however performance was also not consistent between these sessions, $c = 0.174, p=0.386$.

Again DS performed outside of 3 SD's of the control participant's performance (see Figure 42). Modified t-tests confirmed that the difference to controls was significant $F(1, 8) = 38.25, p < 0.001$.

Pictures

DS correctly answered an average of 17/24 (71%), whereas control participants correctly answered an average of 23.11/24 (96.3% ± 0.60) questions correctly (DS only performed this task once, therefore consistency was not measured).

Again DS performed outside of 3 SD's of the control participant's performance (see Figure 42). Modified t-test confirmed that DS's performance was significantly lower than that of the control participants $F(1, 8) = 93.33, p < 0.001$.

Loglinear analysis was performed on DS's data, with the factors being stimuli type (sound vs. picture) and accuracy (correct vs. incorrect). This revealed that DS was more incorrect than correct, $\chi^2 = 4.143, p = 0.042$.

Recognition Task

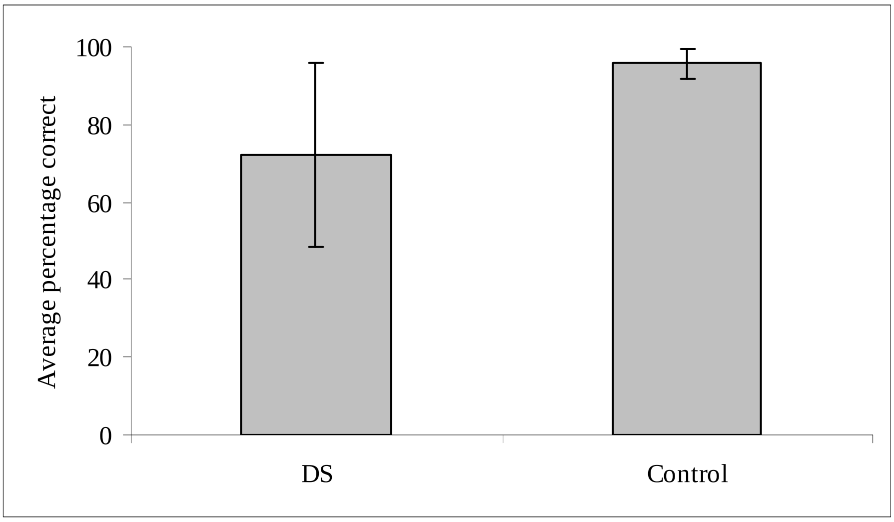


Figure 43: Percentage of correct responses on instrument recognition task

DS correctly matched an average of 26/36 (72.22%) instrument clips to the appropriate pictures. McNemar test revealed that performance was significantly better in the second session (10/18 for session 1, 16/18 for session 2, $p = 0.031$), and performance was not consistent between sessions, $0.368, p=0.094$.

Control participants correctly matched an average of 17.22/18 (94.44%; SD = 0.67). Therefore DS again performed below 3 SD of the mean of the control participants' performance (see Figure 43). However modified t-test revealed that DS's performance was not significantly different from the control participants $F(1, 8) = 36.43$, $p = 0.095$.

DS also completed this task with the sounds and pictures appearing at the same time, rather than separated. He correctly matched 16/18 (88.9%) sounds to pictures. Despite being higher that the average score when the sounds were presented before the picture, this score is actually the same as that achieved in the second session of the latter task.

Speech

Spoken word-picture matching task: DS correctly matched 35/40 (87.5%) words, which was below the level of the 31 control participants tested using the PALPA 39.29/40 (98.3%, SD = 1.07); DS performed below 3 SD's of the mean for the control participants.

Spoken word matching task: DS correctly matched 62/72 (86.1% - same: 31/37; different: 31/37). Norms acquired from 24 normal participants revealed an average performance of 70.4/72 (97.7% - same: 35.5/37; different: 34.9/37, SD = 1.68). DS performed below 3 SD's of the mean for the control participants.

Summary of findings

Table 35: Summary of DS's performance

			Matching							
	Naming		Exact		Two exemp.		Classif.		SRT/ IRT	Tune Recog
	S	P	S	P	S	P	S	P		
Animal	×	×	✓	×	×	×	×	×	×	n/a
Object	×	×	✓	×	×	×	×	×	×	n/a
Music	×	×	×	✓	×	×	×	×	✓	×

×= impaired; ✓= unimpaired; S= sound; P=picture;
SRT = Sound Recognition Task; IRT = Instrument Recognition Task

Table 35 is a summary of DS performance on the animal, object and music tasks, which shows that he was impaired on most tasks, except for the exact matching tasks for animal and object sounds and the recognition task for musical instruments.

Discussion

Music

The same tasks that were carried out with animals and objects were also investigated with musical instruments. DS was impaired relative to control at naming musical instruments, though he could mime more instruments than he could name. In addition in comparison to naming of animals and objects, DS was poorer at naming musical instruments.

His performance on the exact matching tasks revealed that he was impaired compared to control at matching musical instrument sounds but relatively unimpaired at matching pictures of musical instruments. Performance on the two exemplars matching task revealed that DS was impaired at matching both sounds and pictures.

DS was also impaired on both the sound and the picture versions of the classification task, compared to control participants.

This pattern of performance is similar to that of the non-musical sounds, in that he appears to have a stronger associative deficit compared to apperceptive impairments. He was more impaired at tasks involving musical instruments, than non-music items however.

Speech

DS performance on the matching tasks of the PALPA revealed that he was impaired at both matching spoken words to pictures, and matching pairs of spoken words, in comparison to control, suggesting that he has verbal recognition impairments also.

General Discussion

The current study aimed to investigate auditory recognition abilities of a patient with left frontal lobe, to determine how such impairments are related to visual recognition abilities and to investigate how sound category effects recognition. Tasks were used that aimed to investigate different aspects of auditory recognition, in a similar way to studies of visual recognition. A number of tasks were employed to assess recognition abilities (i.e. sound naming, picture naming and sound-picture matching tasks), apperceptive impairments (i.e. exact matching of two stimuli) and associative impairments (i.e. matching of two exemplars of stimuli, sound-picture matching and classification tasks).

Consistency

Overall DS's performance appeared to be consistent across the two sessions of each task, suggesting that the deficits in his performance were due to a semantic impairment, rather than due to a disorder of access to the correct representation (Warrington et al., 1984). In a number of tasks however, DS's performance was not consistent. In such tasks, DS generally had to make a forced choice response with one of two answers and it is possible that on some trials he guessed at the correct answer. This could lead to an inconsistent pattern of responding since an item guessed correctly in one session might be guessed incorrectly in the next. Overall the data fit a pattern of item-consistency.

Recognition abilities

DS showed deficits compared to controls in naming of all sound types; he was most impaired at naming musical instruments, followed by animal sounds and object sounds.

DS was generally better at naming pictures compared to sounds; however he was still impaired, relative to controls. He was particularly impaired at naming pictures of musical instruments and also animal pictures, however remained unimpaired at naming pictures of objects.

This dissociation between animal and object picture naming (and to a smaller extent between animal and object sound naming) appears to suggest a category-specific impairment for animal sounds.

It is possible however that people are less familiar with animal sounds and pictures than object sounds and picture. The number of easily recognisable sounds however is actually quite small (Kraut et al., 2006), therefore a fairly small selection of sounds were used in each task, and the most familiar of these categories were chosen, therefore it is unlikely that the sounds used were truly unfamiliar. A number of studies have found that when the familiarity and frequency of the stimuli are controlled for category-specific impairments disappeared (Stewart et al., 1992; Funnell and Sheridan, 1992; Gaffan and Heywood, 1993, cited in Caramazza and Shelton, 1998). In contrast Caramazza and Shelton (1998) found that category-specific deficits remained, even when familiarity and frequency were controlled for. In the current study, familiarity and frequency were not controlled, however the sounds were rated for familiarity by 11 young control participants, which revealed no significant differences between animal, environmental sounds or music, suggesting that familiarity is not an explanatory factor in DS impairment with animal sounds. Also DS's performance was compared to control participants; therefore presumably any difference in familiarity and frequency between categories would also affect the control participants.

The participants were also asked to match animal sounds, environmental sounds and musical instrument sound clips to pictures of the associated sounds. DS was impaired compared to controls at matching animal and environmental sounds to the relevant pictures; however he made equal amounts of acoustic and semantic errors, suggesting that his deficit is not overly due to either an apperceptive or associative impairment. In contrast DS was relatively unimpaired at matching instrument sounds to pictures of the instruments, despite having the poorest performance at naming items in this category.

DS performance on the recognition tasks (i.e. sound-picture matching) when the sounds appeared simultaneously with the pictures revealed little difference to when the sound presentation was temporally separated from the picture presentation. This suggests that DS does not benefit from using the pictures as cues when the sounds and pictures appear together.

One reason for this difference between animal and environmental sounds, and musical instruments, is that the matching task for animal and environmental sounds had four answer options, (correct answer, semantic foil, acoustic foil and odd item), whereas the musical instrument matching task had only two answer options. Therefore the musical task may have been much easier for DS to do, compared to the sound recognition task.

Alternatively DS's difficulty at naming musical instruments sounds and pictures may indicate an apperceptive impairment for musical instruments, while intact performance at the sound-picture matching may indicate spared associative abilities. His poor performance on the exact matching task for musical instruments supports this as perceptive abilities are needed to perform this task. DS performance on the two exemplars matching task does not fit the proposed pattern of impairments, as if DS has spared associative abilities, he should be unimpaired at this task; however he performed below the level of control participants

Apperceptive Deficits

Apperceptive deficits are difficulties in recognising the perceptual aspects of a stimulus. The exact matching tasks investigated this ability as they required participants to consider the physical aspects of the stimuli in order to accurately determine whether two sounds were exactly the same. DS was unimpaired at matching pairs of animal sounds and environmental sounds, but he was impaired when asked to match musical instrument sounds. When asked to match pictures however, DS was unimpaired at matching both object pictures and musical instrument pictures however he did show impairments in matching animal pictures.

This pattern of findings suggest that DS does not have an apperceptive deficit for environmental sounds or pictures, or animal sounds, but does for animal pictures and music sounds.

DS difficulties with music were investigated further by determining whether he also had difficulties with melody recognition. In this task, DS was asked to indicate whether two tunes were exactly the same or different. DS was also impaired at melody recognition when compared to controls; therefore his impairments with musical stimuli extend across musical instruments to melody also.

Associative Deficits

Associative abilities enable people to match different exemplars of target stimuli. DS was impaired at matching two exemplars of animal sounds, environmental sounds and musical instruments.

DS and the controls were also asked semantic questions about the sound and picture stimuli. This revealed that DS was impaired at answering questions about all sound categories, but was more impaired at answering questions about animal sounds compared to environmental sounds. This dissociation was stronger when asked to complete the same task with pictures. This therefore suggests that DS has a general associative problem for animals, which results in impairments in answering questions about animal sounds and pictures, but also in matching two exemplars of animal sounds.

He was also impaired at answering questions about musical instruments from their sounds and from their pictures.

These tasks suggest that DS has deficits in associating different sound and different picture exemplars and also in associating sounds and pictures to semantic knowledge about these stimuli.

DS completed most of the task on two separate occasions, enabling consistency in his responses to be measured. For all tasks, except the instrument recognition task, there were no significant differences between the two sessions. In contrast contingency coefficients were not significant for musical instrument picture identification, the majority of the matching tasks, the sound classification tasks or the recognition tasks Importantly, the deficits in auditory recognition reported here occurred in the absence of significant hearing impairments, as pure tone audiometry revealed that DS had mild impairment for his age in the higher frequency ranges only.

This study also reveals auditory recognition impairments, in the absence of temporal lobe damage, as is usually required for such deficits. Rather, DS had a large lesion of the left frontal lobe. Few studies have reported auditory recognition disorders in the absence of temporal lobe lesions, however Johkura et al. (1998) reported a patient who had a small lesion in the inferior colliculi, but no lesions of the temporal lobes. This patient was impaired at environmental sound identification and familiar song identification, but unimpaired at tone discrimination and melody discrimination. The current study also found impaired environmental sound identification, but impaired rhythm discrimination also.

A key feature about the current study is the attempt to develop auditory counterparts to common tasks used to assess visual recognition disorders. A number of studies have made attempts to investigate the dissociations between apperceptive and associative auditory processing disorders (Buchtel et al., 1989; Clarke et al., 1996; Vignolo, 1982; Mendez, 2001). However few studies have looked at these dissociations as a function of sound category, or in relation to visual recognition abilities. The current study therefore gives a clearer assessment of possible auditory processing disorder. Further research with patient with temporal lobe lesions, using the current assessment battery is now warranted to assess the existence of auditory apperceptive/associative deficits, auditory category-specific impairments, and the relationship between auditory and visual processing disorders.

Table 36: Animal sounds and pictures used in experimental tasks

Tasks	Item	No. of controls named sounds	No. of controls named pictures
aa, ab, ac, srt	Cat	9	9
ac, srt	Chimp	9	9
aa, ab	Cockerel	8	9
aa, ab, ac, srt	Cow	9	9
aa, ab, ac, srt	Dog	9	9
Ac	Dolphin	9	9
aa, ab, ac	Duck	9	9
Ac	Elephant	9	9
aa, ab, ac, srt	Frog	9	9
Srt	Goat	9	9
aa, ab, srt	Horse	4	9
ac, srt	Lion	9	9
aa, ab	Owl	9	9
Ac	Parrot	9	9
aa, ab, ac	Pig	9	9
aa, ab, ac	Sheep	9	9
Srt	Wasp	9	9

aa = Exact Match; ab = Two Exemplar Match; ac = Classification Task; srt = Sound Recognition Task

Table 37: Object sounds and pictures used in experimental tasks

Tasks	Item	No. of controls named sounds	No. of controls named pictures
aa, ab, srt	Airplane	6	9
Srt	Alarm clock	8	9
aa, ab, srt	Camera	0	9
ac	Car horn	9	9
ac	Chopping	9	9
ac	Cork	9	9
aa, ab	Doorbell	9	9
ac	Drill	9	9
Srt	Gun	9	9
aa, ab, ac	Hammer	9	9
Srt	Heels	9	9
ac, srt	Helicopter	9	9
ac	Motorbike	9	9
aa, ab, ac	Saw	9	9
Srt	Shower	9	9
aa, ab, ac, srt	Sweeping	9	9
aa, ab, ac, srt	Telephone	9	9
aa, ab, ac	Toilet	9	9
ac	Toothbrush	9	9
aa, ab, srt	Train	9	9
aa, ab	Typewriter	9	9
Srt	Wind	9	9

aa = Exact Match; ab = Two Exemplar Match; ac = Classification Task; srt = Sound Recognition Task

Table 38: Percentage of correct responses of DS in each animal and object task and in each session

Task	Stimuli Type	Stimuli Category	Session One score (%)	Session Two score (%)
Naming	**Sound**	**Animal**	23.53	29.41
		Object	33.33	27.27
		Total	28.95	28.21
	Picture	**Animal**	58.82	58.82
		Object	90.91	95.45
		Total	76.92	79.49
Matching: Exact	**Sound**	**Animal**	95.00	100.00
		Object	100.00	100.00
		Total	97.50	100.00
	Picture	**Animal**	70.00	80.00
		Object	90.00	100.00
		Total	80.00	90.00
Matching: Two exemplar	**Sound**	**Animal**	70.00	65.00
		Object	70.00	75.00
		Total	70.00	70.00
	Picture	**Animal**	65.00	60.00
		Object	80.00	80.00
		Total	72.50	70.00
Classification	**Sound**	**Animal**	58.33	58.33
		Object	79.17	70.83
		Total	68.75	64.58
	Picture	**Animal**	58.33	50.00
		Object	83.33	79.17
		Total	70.83	64.58
Recognition		**Animal**	55.56	55.56
		Object	50.00	50.00
		Total	52.63	52.63

Table 39: Musical instruments used in experimental tasks

Task	Instrument	No. of controls named sounds	No. of controls named pictures
ma, mb, irt	accordion	9	9
ma, mc, irt	bagpipes	9	9
ma, mc, irt	cello	9	9
ma, mc, irt	clarinet	9	9
ma, mb, mc irt	drums	9	9
ma, mb, mc, irt	flute	9	9
ma, mb, mc, irt	guitar	9	9
ma, mb, irt	harmonica	9	9
ma, mb, mc, irt	harp	9	9
ma, irt	oboe	9	9
ma, mb, irt	organ	9	9
ma, mb, irt	piano	9	9
ma, mc, irt	saxophone	9	9
ma, mc, irt	tambourine	9	9
ma, mb, mc, irt	trumpet	9	9
ma, irt	tuba	9	9
ma, mb, mc, irt	violin	9	9
ma, mc, irt	xylophone	9	9

ma = Exact Match; mb = Two Exemplar Match; ac = Classification Task; irt = Sound Recognition Task

Table 40: Percentage of correct responses of D.S. in each musical instruments task and in each session

Task	Stimuli Type	Session One (%)	Session Two (%)
	Sound	27.78	22.22
Naming	**Picture**	50.00	72.22
Matching:	**Sound**	77.78	91.67
Exact	**Picture**	88.89	100.00
Matching:	**Sound**	63.64	54.55
Two exemplar	**Picture**	68.18	59.09
	Sound	50.00	66.67
Classification	**Picture**	70.83	not tested
Recognition		55.56	88.89
Tune recognition		50.00	56.67